IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

POLAROID CORPORATION, Plaintiff and Counterclaim Defendant, v. HEWLETT-PACKARD COMPANY, Defendant and Counterclaim Plaintiff.	C.A. No. 06-738-SLR

REDACTED

DECLARATION OF RAYMOND N. SCOTT, JR. IN SUPPORT OF DEFENDANT HEWLETT-PACKARD'S MEMORANDUM IN OPPOSITION TO PLAINTIFF POLAROID CORPORATION'S MOTION FOR SUMMARY JUDGMENT THAT CLAIMS 1-3 OF U.S. PATENT NO. 4,829,381 ARE NOT OBVIOUS

FISH & RICHARDSON P.C.
William J. Marsden, Jr. (#2247)
Raymond N. Scott, Jr. (#4949)
919 N. Market Street, Suite 1100
Wilmington, DE 19801
Tel.: (302) 652-5070
Fax: (302) 652-0607
Email: marsden@fr.com; rscott@fr.com

Robert S. Frank, Jr. *(pro hac vice)*
Carlos Perez-Albuerne *(pro hac vice)*
CHOATE, HALL & STEWART LLP
Two International Place
Boston, MA 02109
Tel.: (617) 248-5000
Fax: (617) 248-4000
Emails: rfrank@choate.com; cperez@choate.com

John E. Giust *(pro hac vice)*
Matthew E. Bernstein *(pro hac vice)*
MINTZ, LEVIN, COHN, FERRIS, GLOVSKY AND POPEO PC
5355 Mira Sorrento Place, Suite 600
San Diego, CA 92121-3039
Tel.: (858) 320-3000
Fax: (858) 320-3001
Emails: jgiust@mintz.com; mbernstein@mintz.com

Attorneys for Defendant and Counterclaim-Plaintiff Hewlett-Packard Company

Dated: June 5, 2008

I, Raymond N. Scott, Jr., declare as follows:

1. I am an attorney with Fish & Richardson P.C., counsel for Hewlett-Packard Company. I am a member of the Bar of the State of Delaware and of this Court. I have personal knowledge of the matters stated in this declaration and would testify truthfully to them if called upon to do so.

2. Attached hereto as Exhibit A is a true and correct copy of excerpts of the Deposition of Dr. Rangaraj Rangayyan.

3. Attached hereto as Exhibit B is a true and correct copy of excerpts of the Deposition of Dr. Peggy Agouris.

4. Attached hereto as Exhibit C is a true and correct copy of U.S. Patent No. 4,528,584 issued to inventor Mohammed S. Sabri.

5. Attached hereto as Exhibit D is a true and correct copy of U.S. Patent No. 4,654,710 issued to inventor Christian J. Richard.

6. Attached hereto as Exhibit E is a true and correct copy of U.S. Patent No. 4,789,933 issued to inventors Chen et al.

7. Attached hereto as Exhibit F is a true and correct copy of excerpts of the text "Digital Image Processing," by Gonzalez R.C. and Wintz P., (Addison-Wesley, Reading, MA, 1977).

8. Attached hereto as Exhibit G is a true and correct copy of excerpts of the article "Digital Image Enhancement and Noise Filtering by Use of Local Statistics," by Jong-Sen Lee, (IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 2, pp. 162-168, March 1980).

9. Attached hereto as Exhibit H is a true and correct copy of excerpts of the article "Real-Time Adaptive Contrast Enhancement," by Patrenahalli M. Narendra and Robert C, Fitch (IEEE Transaction on Pattern Analysis and Machine Intelligence, VOL. PAMI-3, No. 6, pp. 655-661, November 1981).

10. Attached hereto as Exhibit I is a true and correct copy of excerpts of the article, "Digital Image Enhancement: A Survey", by David C. Wang, Anthony H. Vagnucci and C.C. Li, (Computer Vision, Graphics, and Image Processing, Vol. 24, pp 363-381 (1983)).

11. Attached hereto as Exhibit J is a true and correct copy of excerpts of the article, "Feature Enhancement of Film Mammograms using Fixed and Adaptive Neighborhoods," by Gordon R and Rangayyan RM, Applied Optics, 1984, 23(4): 560-564.

I declare under penalty of perjury under the laws of the United States of America that the foregoing is true and correct.

Executed this 5th of June, 2008, at Wilmington, Delaware.

/s/ Raymond N. Scott, Jr.
Raymond N. Scott, Jr.

CERTIFICATE OF SERVICE

I hereby certify that on June 5, 2008, I electronically filed with the Clerk of Court the foregoing document using CM/ECF which will send electronic notification of such filing(s) to the following counsel:

Via Email
Jack B. Blumenfeld (#1014)
Julia Heaney (#3052)
Morris, Nichols, Arsht & Tunnell, LLP
1201 North Market Street
Wilmington, DE 19899-1347
Phone: 302-658-9200
Fax: 302-658-3989
Emails: jblumenfeld@mnat.com; jheaney@mnat.com

Attorneys for Plaintiff and Counterclaim-Defendant Polaroid Corporation

Via Email
Russell E. Levine, P.C.
Michelle W. Skinner/David W. Higer
Maria A. Meginnes/Courtney Holohan/C. Beasley
Kirkland & Ellis LLP
200 East Randolph Drive
Chicago, IL 60601
Phone: 312-861-2000
Fax: 312-861-2200
Emails: rlevine@kirkland.com; ggerst@kirkland.com; mskinner@kirkland.com; dhiger@kirkland.com; mmeginnes@kirkland.com; mmeginnes@kirkland.com; cbeasley@kirkland.com

Attorneys for Plaintiff and Counterclaim-Defendant Polaroid Corporation

Courtesy Copy Via Federal Express
Michelle W. Skinner
Kirkland & Ellis LLP
200 East Randolph Drive
Chicago, IL 60601
Phone: 312-861-2000
Fax: 312-861-2200

/s/ Raymond N. Scott, Jr.
Raymond N. Scott, Jr.

Exhibit A

REDACTED
IN ITS ENTIRETY

Exhibit B

REDACTED
IN ITS ENTIRETY

Exhibit C

United States Patent [19]

Sabri

[11] **Patent Number: 4,528,584**

[45] **Date of Patent: Jul. 9, 1985**

[54] **BILEVEL CODING OF COLOR VIDEO SIGNALS**

[75] Inventor: **Mohammed S. Sabri**, Beaconsfield, Canada

[73] Assignee: **Northern Telecom Limited**, Montreal, Canada

[21] Appl. No.: **446,608**

[22] Filed: **Dec. 3, 1982**

[51] **Int. Cl.**[3] **H04N 9/32**

[52] **U.S. Cl.** **358/12**; 358/13

[58] **Field of Search** 358/12, 13

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,467,346 8/1984 Mori 358/13

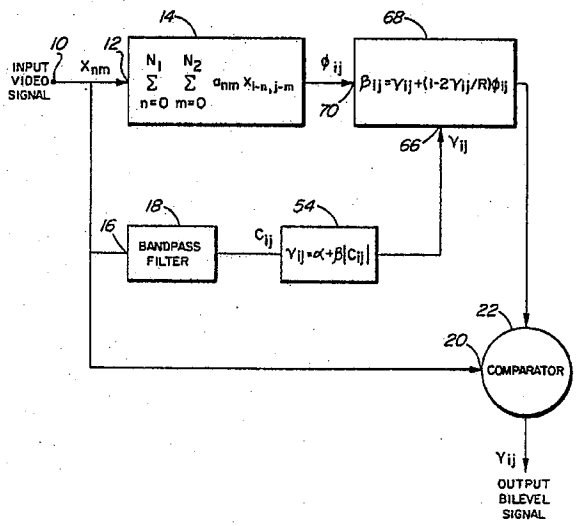
Primary Examiner—Richard Murray

Attorney, Agent, or Firm—Thomas Adams

[57] **ABSTRACT**

Bilevel coding of color video signals, for example video teleconference or NTSC broadcast television signals, is used to reduce the required storage capacity or transmission channel bandwidth. Each multi-level or continuous tone picture element in a frame is compared to a threshold and assigned one of two values depending upon whether or not it exceeds the threshold value. The threshold is produced by averaging the luminance components of neighboring picture elements and constraining the average by a signal derived from the chrominance of the original video signal. A significant feature is that the use of the chrominance signal in this way obviates the need, common to known monochrome systems, of adding noise (or dither signal) to the signal with concomitant picture degrading artifacts.

25 Claims, 2 Drawing Figures



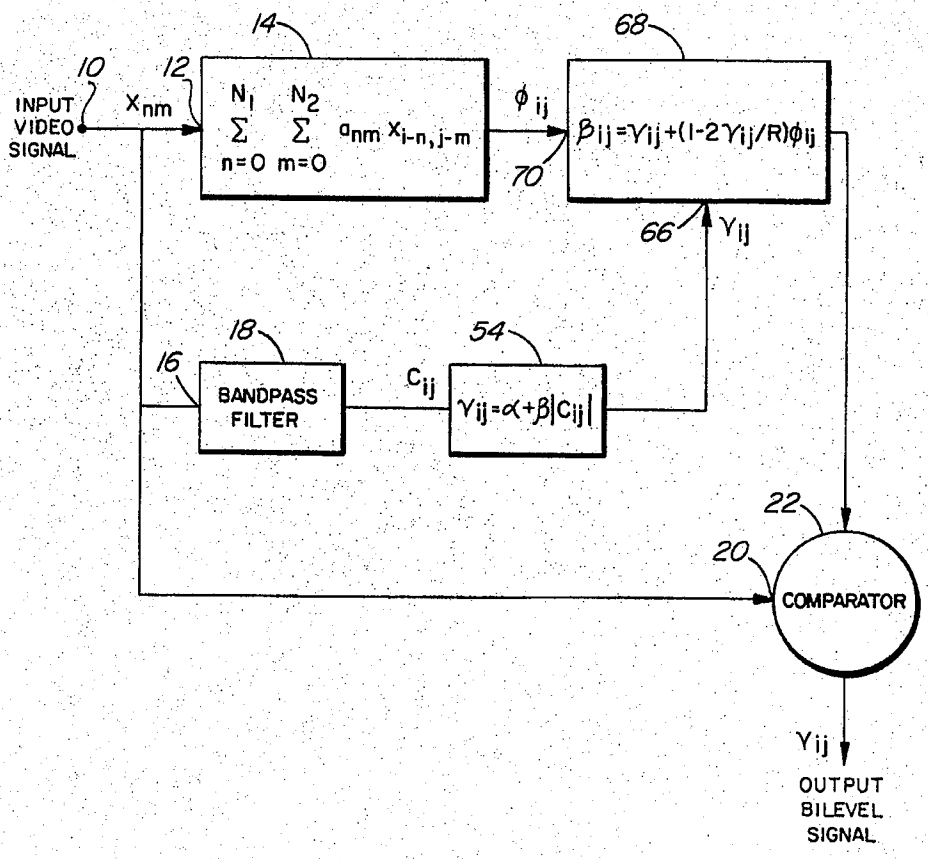

14
68
10
12
INPUT VIDEO SIGNAL
x_{nm}
$\sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} x_{i-n,j-m}$
ϕ_{ij}
$\beta_{ij}=\gamma_{ij}+(1-2\gamma_{ij}/R)\phi_{ij}$
70
66
γ_{ij}
18
16
BANDPASS FILTER
c_{ij}
54
$\gamma_{ij}=\alpha+\beta|c_{ij}|$
22
20
COMPARATOR
γ_{ij}
OUTPUT BILEVEL SIGNAL

FIG. 1

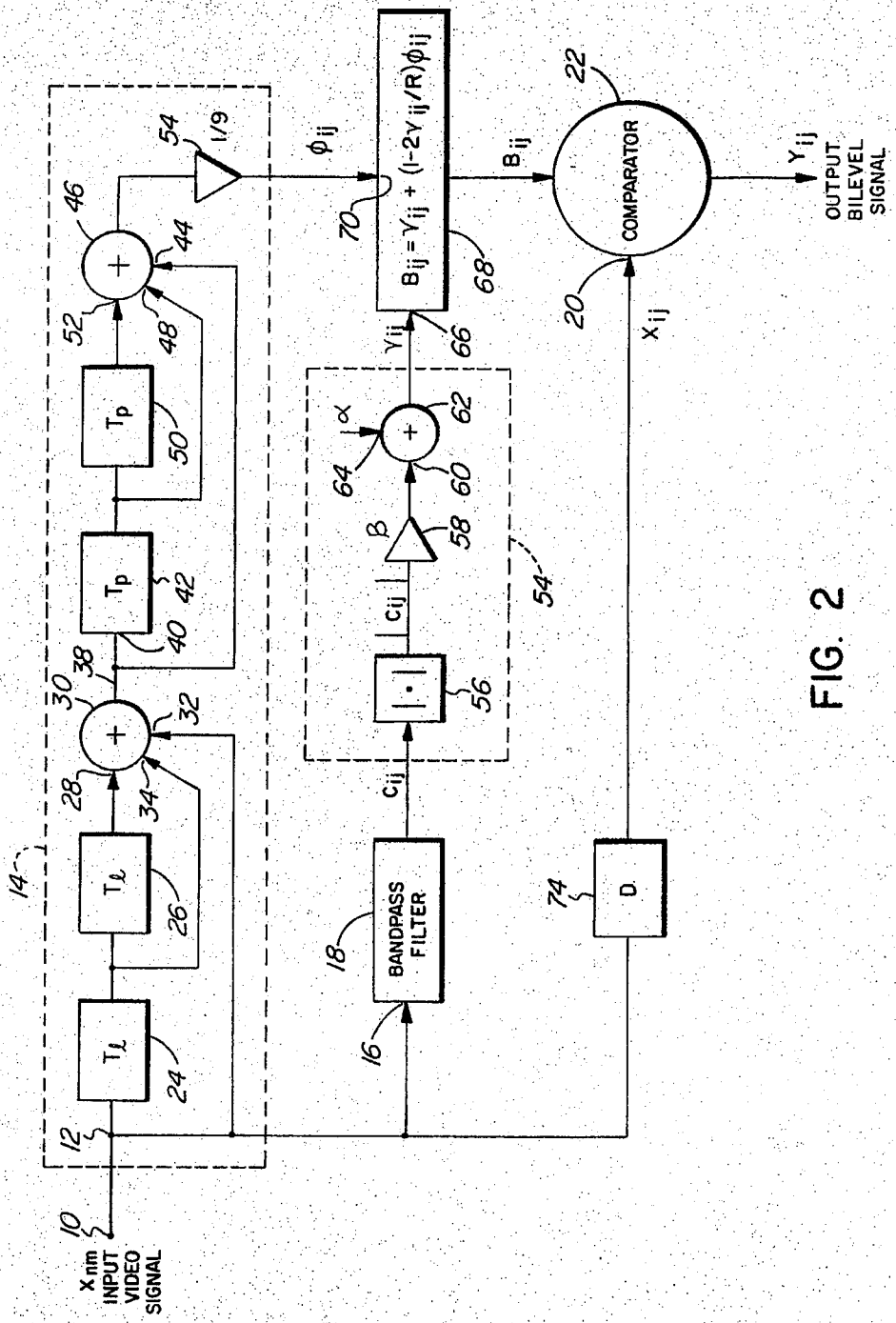

X_{nm} INPUT VIDEO SIGNAL
10
12
14
16
18
20
22
24
26
28
30
32
34
38
40
42
44
46
48
50
52
54
1/9
ϕ_{ij}
70
$B_{ij} = Y_{ij} + (1-2Y_{ij}/R)\phi_{ij}$
68
B_{ij}
COMPARATOR
Y_{ij}
OUTPUT BILEVEL SIGNAL
66
Y_{ij}
X_{ij}
T_ℓ
T_p
T_p
T_ℓ
BANDPASS FILTER
C_{ij}
C_{ij}
α
β
62
64
60
58
56
54
D
74

FIG. 2

BILEVEL CODING OF COLOR VIDEO SIGNALS

The invention relates to video signal processing, particularly to a method and apparatus for bilevel coding or representation of colour video signals, and is especially applicable to composite television signals, for example NTSC broadcast colour television signals.

Whether a video signal originates as a broadcast television signal or otherwise, for example in an interactive visual communications system using video terminals and telephone lines, the need often arises to code the signal to reduce the required storage capacity or transmission channel bandwidth. Also in television programme production it is sometimes desirable to create a stylized representation or a sketch of the picture for special effects or animation. A type of process evolved for this purpose is so-called "bilevel coding (representation)". In such a process, each multi-level or continuous tone point in the frame, or picture element (Pel), is represented in bilevel form, i.e. it can assume only one of two states, typically high or low. Consequently only one digital bit is required to represent each picture element and the receiver can be very simple.

Basically bilevel coding involves comparing the picture element intensity with a threshold value and, depending upon whether or not it exceeds the threshold, assigning a corresponding output signal to either a high or low state. Generally the threshold value is chosen, and hence varied, according to the content of an area of the picture surrounding or adjacent to the picture element being coded at a particular instant. Determination of the threshold value affects the ability to preserve grey scale rendition, apparent contrast and detail in the picture.

Hitherto bilevel coding has generally been limited to monochrome signals. One known process, known as the "ordered dither" technique is disclosed in a paper entitled "Design of dither waveforms for quantized visual signals", by J. O. Limb. Bell System Technical Journal No. 48, 1969 pp. 2555–2582 and a paper entitled "An optimum method for two level rendition of continuous tone pictures", by B. E. Bayer, ICC conference record, 1973 pp. 26–11, 26–15.

In this process a set of predetermined position-dependent thresholds are used to derive the bilevel picture elements or a random dither signal, for example noise, is added to the multilevel picture. The process is primarily suited to bilevel displays since an objectionable flicker is usually perceived if the signal is displayed on a television monitor. Also, the process is not satisfactory for colour television signals due to excessive high frequency components in the chrominance bands which overlap with the chrominance signal and distort it.

An alternative process, known as the "constrained average technique" is disclosed in U.S. Pat. No. 3,961,134 issued June 1, 1976, which is hereby incorporated by reference. In this process the threshold value is chosen as a function of the local average within an area surrounding the point or picture element. To produce an apparent grey scale the process utilizes noise in the video signal. To alleviate the reliance upon an indeterminate parameter a controlled amount of noise (dither) is added to the video signal, in effect utilizing the ordered dither technique. Accordingly, the resulting picture suffers from the same problems of flicker and chrominance band distortion.

An object of the present invention is to eliminate or at least mitigate these problems in providing a process and apparatus for bilevel coding of a colour video signal.

According to one aspect of the present invention a process for bilevel coding of a colour video signal includes the steps of:

(i) deriving from a plurality of picture elements a first signal proportional to the luminance component of the colour video signal;

(ii) deriving a second signal proportional to the chrominance component of the colour video signal;

(iii) computing from the first and second signals a threshold value;

(iv) comparing a picture element value to the threshold value and deriving an output signal having one of two states in dependence upon whether or not the picture element value exceeds the threshold value.

In preferred embodiments of the invention the signal proportional to the luminance component is a weighted constrained average (ϕ_{ij}), computed, for example, in accordance with the equation

$$\phi_{ij} = \sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} X_{i-n,j-m}$$

where X is the pel value and a_{nm} is the weighting coefficient chosen to the signal proportional to the luminance.

Then the chrominance-proportional signal (C_{ij}), which acts as a dither signal producing an apparent grey scale, is used to derive a contrast enhancement factor X_{ij} thus:

$$\gamma_{ij}=\alpha+\beta|C_{ij}|$$

where α and β are constants. Conveniently α is in the range 5–20, preferably 10, and β is in the range 0.5–2, preferably 1.0. $|C_{ij}|$ denotes the absolute value or magnitude of C_{ij}.

The threshold value B_{ij} is then computed as:

$$B_{ij}=\gamma_{ij}+(1-2\gamma_{ij}/R)\phi_{ij}$$

Where γ_{ij} is the contrast enhancement factor and R is the maximum range of the video signal, for example 256 for an 8 bit digital signal.

According to another aspect, the invention comprises apparatus for implementing the process of the first aspect. Thus, said apparatus comprises:

means for deriving from a plurality of picture elements a first signal proportional to the luminance component of the colour video signal;

means for deriving a second signal proportional to the chrominance component of the colour video signal;

means for computing from the first and second signals a threshold value; and

means for comparing a picture element value to said threshold value to derive an output signal having one of two states in dependence upon whether the picture element value exceeds the threshold value or not.

Means for deriving the chrominance-proportional signal may conveniently comprise a bandpass filter having a pass band encompassing the chrominance subcarrier frequency f_{sc}. Such bandpass filter may be a one- or two-dimensional digital filter.

A particularly simple apparatus may be realized if the signal proportional to the luminance component is a weighted constrained ϕ_{ij} average of the form:

$$\phi_{ij} = \sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} X_{i-n,j-m}$$

and furthermore if the weighting coefficient a_{nm} is substantially unity and N_1 is an integer multiple of the subcarrier cycles. Thus if a sampling frequency of $3f_{sc}$ were used, N_1 would be multiples of three.

An exemplary embodiment of the invention will now be described with reference to the accompanying drawings, in which:

FIG. 1 is a schematic block representation of apparatus for bilevel coding of a colour video signal; and

FIG. 2 is a more detailed representation of the apparatus for a sampling frequency triple the colour subcarrier frequency.

It should be understood that although the input video signal can be analog or digital, for the specific embodiment the input signal is in digital form, for example 8 bits, and could be derived from a store, wherein it is stored still in digital form, or could be derived from a camera or other such source with an intervening analog/digital converter to convert it to the digital format required for input to the apparatus. The output bilevel signal will be decoded in the receiver. Since the receiver decoding circuit is known, it too is not shown in the drawing. Typically the receiver will comprise a standard monitor with a one bit digital/analog converter.

Referring to FIG. 1, apparatus for bilevel coding an NTSC composite colour video signal comprises an input terminal 10 to which is applied the digital video signal to be encoded sampled at three times the colour subcarrier frequency f_{sc}. The input terminal 10 is connected to, respectively, an input 12 of summing means 14, an input 16 of a bandpass filter 18 and an input 20 of a comparator 22. The summing means 14 serves to compute from the input picture elements X_{ij} a constrained weighted average ϕ_{ij} according to the formula:

$$\phi_{ij} = \sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} X_{i-n,j-m}$$

$X_{i-n,j-m}$ is the picture element preceding the particular element being encoded. As can be seen from FIG. 2, summing means 14 is adapted to sum nine preceding picture elements to produce the weighted constrained average ϕ_{ij}. Thus, summing means 14 comprises two delays 24 and 26, each equal to the duration of one line scan of the signal. The delays 24 and 26 are connected in series between the input 12 and one input 28 of an adder 30. The input 12 also is connected directly to an input 32 of the adder 30. A third input 34 of the adder 30 is connected to the output 36 of the first delay 24. The output of the adder 30, on line 38, is applied to the input 40 of a delay 42, equal to the duration of one picture element (Pel). Line 38 also is connected to an input 44 of a second adder 46. The output of the Pel delay 42 is applied to a second input 48 of the adder 46 and to a second Pel delay 50, the output of which is applied to a third input 52 of the adder 46. The output of adder 46 is connected to a divider 54, where the summed signal is divided by nine to give the signal ϕ_{ij} proportional to the luminance of the composite colour video signal X_{ij}. For this case the individual coefficients a_{nm} for each delay or Pel is equal to one ninth. Sampling at $3f_{sc}$, the sequence of values of each three successive picture elements will be Y+I,

$$Y - \frac{I}{2} + \frac{\sqrt{3}}{2} Q \text{ and } Y - \frac{I}{2} - \frac{\sqrt{3}}{2} Q.$$

where Y is the luminance component and I and Q are the chrominance components. It will be seen that when these are summed the resultant is proportional to luminance only.

As will be described more fully hereafter, other proportions can be used depending upon sampling frequency and number of samples.

Bandpass filter 18 is a one-or two-dimensional digital filter arranged to pass the chrominance signal, C_{ij}. The centre frequency of the passband is f_{sc}, where f_{sc} is the subcarrier frequency. The output of the filter 18 is connected to the input of means 54 for deriving a contrast enhancement factor proportional to the absolute value of C_{ij} plus a constant α. Means 54 comprises a device 56 for deriving the magnitude, (in effect a rectifier), the output of which is applied by way of an multiplier 58, having a multiplication factor β, to an input 60 of an adder 62. A second input 64 of the adder 62 is connected to a reference source having a value α. The output (γ_{ij}) of the adder 62 is connected to an input 66 of means 68 for computing the threshold value B_{ij} according to the equation:

$$B_{ij} = \gamma_{ij} + (1 - 2\gamma_{ij}/R)\phi_{ij}$$

Where γ_{ij} is the contrast enhancement factor and R is the maximum range of the video signal, for example 256 for an 8 bit digital signal.

A second input 70 of means 68 is connected to the output of divider 54. The output of computing means 68 is applied to a second input 72 of comparator 22. The input signal from terminal 12 is applied to the first input 20 of the comparator by way of a delay 74, equal to the delay of the bandpass filter 18. The output of the comparator 22 constitutes the output bilevel signal for storage and/or transmission.

As mentioned previously, the means 14 derives the weighted constrained average ϕ_{ij} from nine picture elements in a 3×3 matrix. The last element, in a conventional scan at the corner of the matrix, is the element X_{ij} applied also to the comparator 22 and hence compared with the threshold value B_{ij}. The output bilevel signal Y_{ij} will be high (S_1) if the element intensity X_{ij} exceeds B_{ij} and low (S_0) if it does not.

As mentioned previously, the values of a_{nm} can be chosen to render a signal proportional to luminance only, even though it has other than nine samples and a different sampling frequency. For example, at four times the colour subcarrier sampling frequency ($4f_{sc}$) the coefficients for a_{nm} can be:

$$\frac{1}{16}\begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ 2 & 0 & 12 & 0 & 2 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix}$$

It should be appreciated that although the specific description is for a digital video signal, the invention is not limited to digital signals but also comprehends analog signal processing. Then means for deriving the luminance-related signal could be a low pass filter, means for

deriving the contrast enhancement factor could be a rectifier, amplifier and adder or summing amplifier, and means for computing the threshold value B_{ij} would use suitable analog devices.

It will also be appreciated that for some applications, such as stylized single frame images, the invention need not be embodied using the precise circuit element described hereinbefore, but could be put into practice using a microprocessor.

With suitable modification the invention could be applied to component colour video signals, i.e. in which the luminance and chrominance components have previously been separated.

A significant advantage of using the chrominance signal to constrain the averaged luminance signal is that the video signal reproduced can be kept relatively free from objectionable artifacts such as flicker or distorted colour. It will be appreciated that to add noise as in the various dither techniques mentioned in the introduction will generally degrade picture quality.

What is claimed is:

1. Video signal processing apparatus comprising means for bilevel coding discrete picture elements of a colour video signal having a luminance component and a chrominance component, including:

means for deriving from a plurality of picture elements including at least one picture element other than the picture element to be encoded a first signal, (ϕ_{ij}), proportional to said luminance component;

means for deriving a second signal (C_{ij}) proportional to said chrominance component;

means for computing from said the first and second signals a threshold signal (B_{ij}) and

a comparator for comparing said picture element to be encoded with said threshold value and providing an output signal Y_{ij} having either one of two states in dependence upon whether or not said picture element exceeds said threshold value.

2. Video signal processing apparatus as defined in claim 1, wherein said means for deriving a luminance-proportional signal comprises means for computing a weighted average (ϕ_{ij}) of the luminance and chrominance components of a plurality of picture elements.

3. Video signal processing apparatus as defined in claim 2, wherein said means for computing is adapted to compute said weighted average ϕ_{ij} in accordance with the formula

$$\phi_{ij} = \sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} X_{i-n,j-m}$$

where a_{nm} is a weighting factor and $X_{i-n,j-m}$ adjacent picture elements' values.

4. Video signal processing apparatus as defined in claim 2, wherein said means for deriving said luminance-proportional signal is arranged to compute said signal from a plurality of neighbouring picture elements.

5. Video signal processing apparatus as defined in claim 4, wherein said plurality of neighbouring picture elements comprise a 3×3 matrix including said picture element to be encoded.

6. Video signal processing apparatus as defined in claim 5, wherein said means for deriving said luminance-proportional signal comprises two line delays connected serially between the input and an adder, the input being connected directly to one adder input, the output of the first delay connected to a second adder input and the second line delay connected to a third adder input.

7. Video signal processing apparatus as defined in claim 6, further comprising two picture element delays connected serially between the output of said adder and a first input of a second adder, the output of the first adder being connected to a second input of the second adder and the output of the first picture element delay being connected to a third input of the second adder.

8. Video signal processing apparatus as defined in claim 1, wherein said means for deriving a chrominance-proportional signal comprises a bandpass filter having a pass band to pass the chrominance signal.

9. Video signal processing apparatus as defined in claim 8, wherein said filter comprises a one- or two-dimensional digital filter.

10. Video signal processing apparatus as defined in claim 8 wherein said means for deriving a chrominance-proportional signal further comprises means responsive to the output of the filter for providing a contrast enhancement factor γ_{ij}.

11. Video signal processing apparatus as defined in claim 10, wherein said means for providing said contrast enhancement factor is arranged to do so in accordance with the formula $\gamma_{ij}=\alpha+\beta|C_{ij}|$ where $|C_{ij}|$ is the magnitude of C_{ij} and α and β are constants.

12. Video signal processing apparatus as defined in claim 11, wherein α is in the range 5 to 20.

13. Video signal processing apparatus as defined in claim 12, wherein α is equal to 10.

14. Video signal processing apparatus as defined in claim 11, wherein β is in the range 0.5 to 2.0.

15. Video signal processing apparatus as defined in claim 14, wherein β is equal to 1.

16. Video signal processing apparatus as defined in claim 1, wherein said means for computing said threshold signal B_{ij} is arranged to do so in accordance with the formula $B_{ij}=\gamma_{ij}+(1-2\ \gamma_{ij}/R)\phi_{ij}$ where γ_{ij} is a contrast enhancement factor related to said chrominance-proportional signal and R is the maximum range of the video signal.

17. A process for bilevel coding of a colour video signal having a luminance component and a chrominance component comprising the steps of:

(i) deriving from a plurality of picture elements including one or more picture elements other than the picture element to be encoded a first signal (ϕ_{ij}) proportional to said luminance component of the video signal;

(ii) deriving a second signal (C_{ij}) proportional to said chrominance component;

(iii) computing from said first and second signals (ϕ_{ij}) and (C_{ij}) a threshold value;

(iv) comparing said picture element to be encoded with said threshold value and providing an output having either one of two states in dependence upon whether or not said picture element exceeds said threshold value.

18. A process as defined in claim 17, wherein the step of deriving a luminance-proportional signal includes computing a weighted average of a plurality of picture elements.

19. A process as defined in claim 18, wherein said weighted average ϕ_{ij} is computed in accordance with the relationship

$$\phi_{ij} = \sum_{n=0}^{N_1} \sum_{m=0}^{N_2} a_{nm} X_{i-n,j-m}$$

where a_{nm} is a weighting coefficient and $X_{i-n,j-m}$ the values of said plurality of picture elements.

20. A process as defined in claim **19**, wherein said plurality of picture elements comprises with the elements to be encoded a 3 by 3 matrix.

21. A process as defined in claim **19**, wherein said threshold value B_{ij} is computed in accordance with the relationship $B_{ij} = \gamma_{ij} + (1\text{-}2\gamma_{ij}/R)\ \phi_{ij}$ where R is the maximum range of the video signal.

22. A process as defined in claim **17**, wherein said signal C_{ij} is derived by extracting the chrominance signal by means of a bandpass filter whose centre frequency is at f_{sc}.

23. A process as defined in claim **22**, wherein a contrast enhancement factor γ_{ij} is computed from the signal C_{ij} in accordance with the relationshp $\gamma_{ij} = \alpha + \beta |C_{ij}|$ where α is in the range 5 to 10, β is in the range 0.5 to 2.0, and $|C_{ij}|$ is the absolute value of C_{ij}.

24. A process as defined in claim **23**, wherein α is equal to 10.

25. A process as defined in claim **22**, wherein β is equal to unity.

* * * * *

Exhibit D

United States Patent [19]

Richard

[11] **Patent Number: 4,654,710**

[45] **Date of Patent: Mar. 31, 1987**

[54] **CONTRAST AMPLIFIER FOR VIDEO IMAGES**

[75] Inventor: **Christian J. Richard**, Noyal Sur Vilaine, France

[73] Assignee: **Thomson CSF**, Paris, France

[21] Appl. No.: **815,932**

[22] Filed: **Jan. 3, 1986**

[30] **Foreign Application Priority Data**

Apr. 1, 1985 [FR] France 85 00105

[51] **Int. Cl.**[4] **H04N 5/57**

[52] **U.S. Cl.** **358/169;** 358/39; 358/160; 364/734

[58] **Field of Search** 358/169, 39, 160, 32, 358/164, 184, 166, 139; 364/734, 715

[56] **References Cited**

U.S. PATENT DOCUMENTS

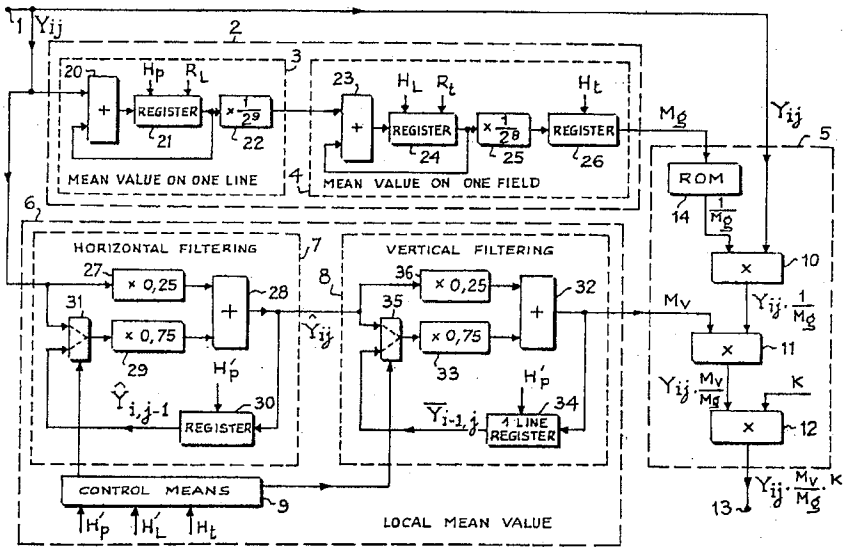
3,638,001	1/1972	Gordon	364/734 X
4,231,065	10/1980	Fitch et al.	358/166
4,553,619	11/1985	Fujinaga	364/734 X
4,603,353	7/1986	Henson	358/39 X
4,606,009	8/1986	Wiesmann	364/734 X

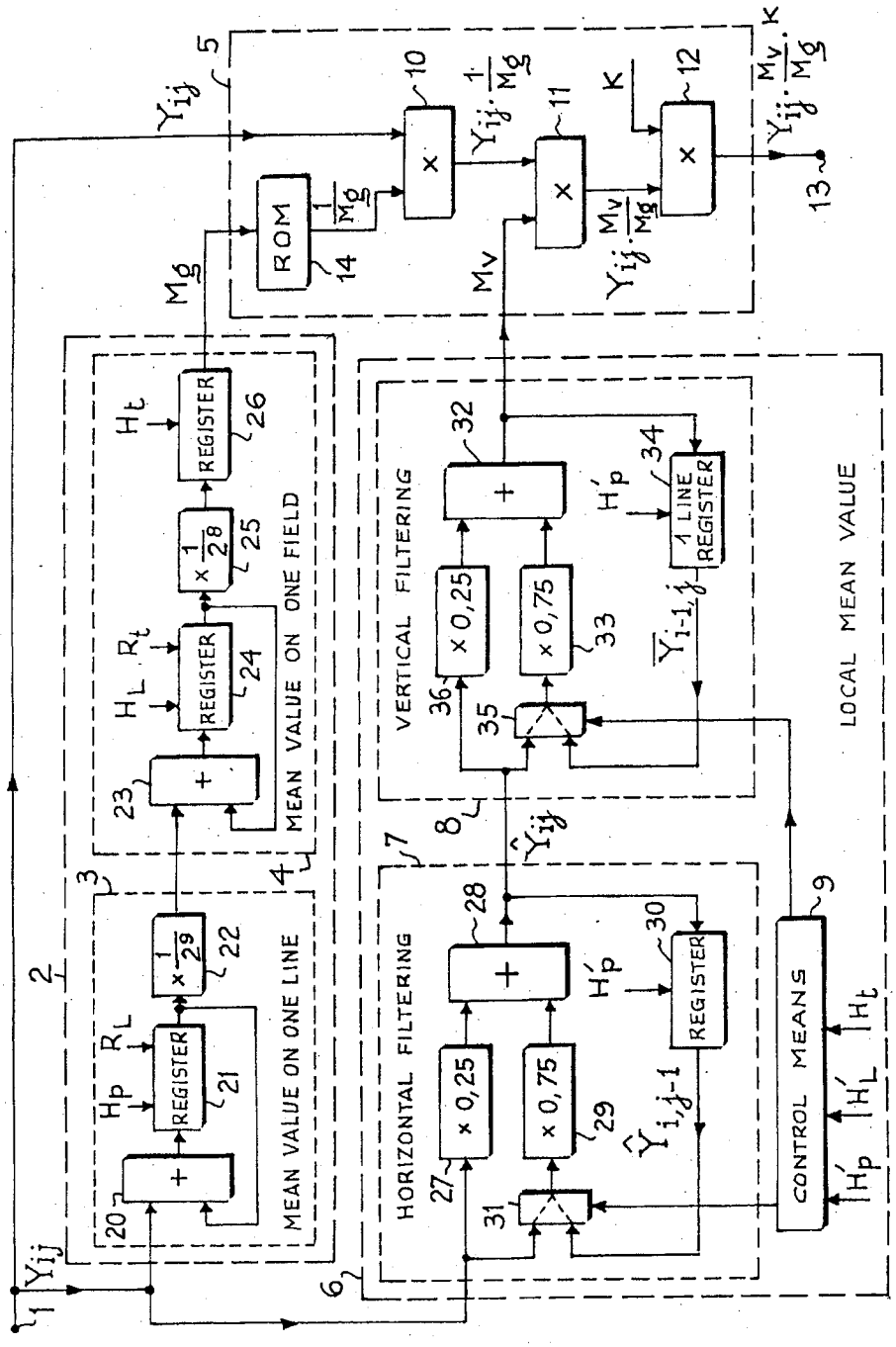
Primary Examiner—Michael A. Masinick
Assistant Examiner—E. Anne Toth
Attorney, Agent, or Firm—Cushman, Darby & Cushman

[57] **ABSTRACT**

A contrast amplifier for video images comprises means for computing a mean value M_g of luminance of all the points of a field, means consisting of a horizontal filter and a vertical filter for computing a local mean value M_v of luminance in the vicinity of a point being processed and means for multiplying the value of luminance of the point during processing by a variable coefficient which is proportional to the ratio M_v/M_g.

3 Claims, 1 Drawing Figure




MEAN VALUE ON ONE LINE
MEAN VALUE ON ONE FIELD
HORIZONTAL FILTERING
VERTICAL FILTERING
LOCAL MEAN VALUE
CONTROL MEANS
ROM
REGISTER
1 LINE REGISTER
x 0,25
x 0,75

CONTRAST AMPLIFIER FOR VIDEO IMAGES

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to devices which are known as contrast amplifiers and are employed for improving the visual perception of video images when these images have low contrast for such reasons as insufficient illumination, for example.

2. Description of the Prior Art

It is a known practice to enhance or amplify the contrast of video images by increasing the gain of the channel used for transmitting the luminance signal representing these images. This increase in gain is either uniform for all values of luminance or variable as a function of the luminance value. The disadvantage of this method is that it also increases the contrast of noise interference. Interference of this type becomes more troublesome as the contrast amplification is greater.

Another known method consists in equalizing the distribution of luminance values of each field throughout the range of possible values. To this end, the distribution of luminance values of a field is determined and each luminance value of the following field is corrected as a function of the distribution which is found in the case of the luminance values of the preceding field. French patent Application No. 2,456,448 filed by the present Applicant describes a method of this type. This known method has the advantage of putting the scale of luminance values to the best possible use but has the effect of increasing noise interference and is complex to apply in practice.

The contrast amplifier in accordance with the invention does not have the disadvantage of enhancing noise interference and is at the same time simple to construct. It is therefore particularly advantageous for processing images which are highly affected by noise. The amplifier makes advantageous use of the fact that noise affects only isolated points or a small number of interrelated points. The amplifier in accordance with the invention does not amplify the visibility of these points since it controls the gain of the channel which transmits the luminance signal, in a different manner at each image point, as a function of the luminance in the vicinity of this point and of the luminance of the entire image. The luminance in the vicinity of each point is determined by means of a so-called bidirectional spatial filtering process which consists of two recursive digital filtering operations performed successively. The mean luminance of an image can be estimated by considering the mean luminance of the preceding image in the case of a sequence of video images.

SUMMARY OF THE INVENTION

In accordance with the invention, a contrast amplifier for video images, each video image being represented by a sequence of digital values of luminance of the points of said image, comprises:

means for estimating a mean value M_g of luminance of all the points of each image in succession;

means for computing a local mean value M_v of luminance in the vicinity of a point being processed;

means for multiplying the value of luminance of the point being processed, by a variable coefficient which is proportional to the ratio M_v/M_g, the value thus obtained being such as to constitute the value of luminance of a point having enhanced contrast.

BRIEF DESCRIPTION OF THE DRAWINGS

Other features of the invention will be more apparent upon consideration of the following description and accompanying drawings in which the single FIGURE is a block diagram of an embodiment of the contrast amplifier according to the invention, processing a sequence of conventional video images. Each video image is composed of two interlaced fields which are processed as two independent images.

DESCRIPTION OF A PREFERRED EMBODIMENT

The embodiment considered by way of example comprises an input terminal **1**, means **2** for computing a mean value M_g of luminance of all the points of a field, means **6** for computing a local mean value M_v of luminance of points which are adjacent to a point being processed for contrast enhancement, means **5** for multiplying the value of luminance of said point by a variable coefficient which is proportional to the ratio M_v/M_g and an output terminal **13** for delivering a sequence of numerical values of luminance having enhanced contrast.

The input terminal **1** receives a sequence of numerical values Y_{ij} of luminance of the points of a field and transmits them to an input of the means **2**, to an input of the means **6** and to a first input of the means **5**. The means **2** has an output connected to a second input of the means **5** for delivering the mean value M_g to said means **5**. The means **6** have an output connected to a third input of the means **5** for delivering the mean value M_v to these latter.

The means **2** are constituted by a first computing device **3** for computing a mean value of luminance on each line and by a second computing device **4** for computing a mean value of luminance on a field by computing the average of the values supplied by the first device **3** during one field. The computing device **3** has an input which constitutes the input of the means **2** and has an output connected to an input of the second computing device **4**. An output of the second computing device **4** constitutes the output of the means **2**.

The first computing device **3** comprises an adder **20**, a single-stage register **21** and a device **22** for dividing by 2^9. The input of the device **3** is connected to a first input of the adder **20**. An output of the adder **20** is connected to a data input of the register **21**. An output of the register **21** is connected to a second input of the adder **20** and to an input of the device **22**. The adder **20** and the register **21** constitute an accumulator. The register **21** receives a clock signal H_p at the point-scanning frequency and receives a zero-reset signal R_L at the line-scanning frequency. The device **22** for dividing by 2^9 is constituted by special wiring for shifting the bits delivered by the register **21** by nine positions. These shifted bits are transmitted to the input of the second computing device **4**.

The second computing device **4** comprises an adder **23**, a single-stage register **24**, a device **25** for dividing by 2^8 and a single-stage register **26**. The input of the computing device **4** is connected to a first input of the adder **23**. An output of the adder **23** is connected to a data input of the register **24**. An output of the register **24** is connected to a second input of the adder **23** and to an input of the device **25**. An output of the device **25** is connected to a data input of the register **26** and an out-

put of this latter constitutes the output of the device **4** and the output of the means **2**.

The adder **23** and the register **24** constitute an accumulator. The register **24** receives a clock signal H_L at the line frequency and receives a zero-reset signal R_t at the field frequency. The device **25** for dividing by 2^8 is constituted by a special wiring which has the effect of shifting the bits delivered by the register **24** by eight positions before transmitting them to the input of the register **26**. The register **26** has a clock input for receiving a clock signal H_t at the field frequency in order to store a mean value of luminance during an entire field period. The mean luminance value M_g computed during a field period is employed for amplifying (enhancing) the contrast during the following field period.

The means **6** for computing a local mean value M_v of luminance in the vicinity of a point being processed are constituted by a horizontal-filtering device **7** connected in series with a vertical-filtering device **8** and by control means **9** for initialization. The filtering devices **7** and **8** are recursive digital filtering devices. The so-called bidirectional spatial filtering achieved by a combination of these devices is comparable with the filtering which would be obtained by computing the arithmetical mean of the luminance values in a sliding "window" centered on the current point. The advantage of this combination, however, lies in the fact that it is conducive to greater simplicity of computations since the horizontal filtering and vertical filtering operations mentioned above are performed by means of two recursive computations which require a smaller number of calculations and less storage of values.

Computation of a mean value of luminance for each line is performed on **512** points, that is to say on a number of points slightly smaller than the number of real points in a line. The number **512** makes it possible to simplify the device **3** since it is very easy to perform a division by 2^9. Computation of the mean value of luminance of the points of a field is carried out on 256 lines, which is smaller than the real number of lines in a field according to European standards since it is an easy matter to perform a division by 2^8. The signal generators for delivering the clock signals H_t, R_L, R_t, H_t are not shown in the FIGURE since their construction is within the capacity of any one versed in the art. The clock signal H_p delivers only 512 pulses per line and the clock signal H_L delivers only 256 pulses per field.

The horizontal filtering device **7** computes, in respect of the j^{th} point on the i^{th} line, a horizontally-filtered luminance value $\hat{Y}_{ij}$ in accordance with the formula:

$$\hat{Y}_{ij}=0.25.Y_{ij}+0.75.\hat{Y}_{i,j-1}$$

where $\hat{Y}_{i,j-1}$ is a value filtered by the filtering device **7** and computed in respect of the immediately preceding point.

Said device **7** comprises a device **27** for multiplication by 0.25, a device **29** for multiplication by 0.75, an adder **28**, a single-stage register **30** and a multiplexer **31**. The multiplication devices **27** and **29** are constituted by read-only memories (ROMs). The input of the means **6** is connected to an input of the device **7** and this latter is connected to an input of the device **27** for multiplication by 0.25 and to a first input of the multiplexer **31**. A second input of the multiplexer **31** is connected to an output of the register **30** and an output of the multiplexer **31** is connected to an input of the device **29** for multiplication by 0.75. A control input of the multiplexer **31** is connected to an output of the control means **9** for initialization.

An output of the device **27** is connected to a first input of the adder **28**. An output of the adder **28** is connected to the output of the device **7** and to an input of the register **30**. An output of the register **30** delivers the value $\hat{Y}_{i,j-1}$ to the second input of the multiplexer **31**. An output of the device **29** is connected to a second input of the adder **28**. The register **30** has a clock input for receiving a clock signal H'_p at the point-scanning frequency and the number of pulses of which corresponds to the real number of points in each line. The generator for the clock signal H'_p is not shown in the FIGURE and construction of this generator is within the capacity of those versed in the art.

The vertical-filtering device **8** has a structure which is similar to that of the horizontalfiltering device **7** but comprises a register **34** for producing a delay which corresponds to one line instead of the register **30** which produces a delay corresponding to one image point. The register **34** delivers a luminance value $\hat{Y}_{i-1,j}$ corresponding to the point equivalent to the point being processed and located on the immediately preceding line, this value being filtered by the devices **7** and **8**. The device **8** comprises a device **36** for multiplication by 0.25, a device **33** for multiplication by 0.75, an adder **32**, the register **34** and a multiplexer **35**.

An input of the device **8** is connected to the output of the device **7** and is connected to an input of the device **36** for multiplication by 0.25 and to a first input of the multiplexer **35**. A second input of the multiplexer **35** is connected to the output of the register **34**. An output of the multiplexer **35** is connected to an input of the device **33** for multiplication by 0.75. A control input of the multiplexer **35** is connected to an output of the means **9**.

An output of the device **36** is connected to a first input of the adder **32**. The output of the adder **32** is connected to the output of the device **8**, this output being such as to constitute the output of the means **6**, and to an input of the register **34**. The output of the register **34** delivers the value $\hat{Y}_{i-1,j}$ to the second input of the multiplexer **35**. An output of the device **33** is connected to a second input of the adder **32**. The register **34** has a control input for receiving the clock signal H'_p.

Since the horizontal-filtering device **7** and the vertical-filtering device **8** are connected in series, they perform a spatial filtering operation which produces a filtered value M_v. This value is similar to the mean value which would be obtained by computing the arithmetical mean of the luminance values of the points located in a sliding window centered on the point being processed. The recursive filtering operation carried out by the device **7** is not possible for the first point of each line and the vertical recursive filtering operation performed by the device **8** is not possible for the first line. It is for this reason that means **9** have been provided for initialization.

The means **9** receive the clock signals H'_p and H_t and a clock signal H'_L. The clock signal H'_L delivers at the line-scanning frequency a number of pulses equal to the real number of lines of each field.

During the period which corresponds to processing of the first point of each line, the means **9** deliver a signal for controlling the multiplexer **31** in such a manner as to ensure that this latter connects the input of the device **29** to the input of the device **7**. Thus the filtered value $\hat{Y}_{i,j-1}$ of luminance of the non-existent preceding

point is replaced by the non-filtered value Y_{ij} of luminance of the point being processed. The device **7** therefore computes: $0.25.Y_{ij}+0.75.Y_{ij}=Y_{ij}$. In the case of the first point of each line, the filtered value is therefore strictly identical with the original value. During each period corresponding to processing of the other points of each line, the means **9** deliver a signal for controlling the multiplexer **31** in such a manner as to ensure that this latter connects the input of the device **29** to the output of the register **30**. The device **7** then computes a filtered value in accordance with the recurrence formula :

$$\hat{Y}_{ij}=0.25.Y_{ij}+0.75.\hat{Y}_{i,j-1}$$

During the period which corresponds to processing of all the points of the first line of each field, the means **9** deliver a signal for controlling the multiplexer **35** in such a manner as to ensure that this latter connects the input of the device **33** to the input of the device **8**. Thus the filtered value $\overline{Y}_{i-1,j}$ of the point equivalent to the current point on the non-existent preceding line is replaced by the horizontally filtered value $\hat{Y}_{ij}$ of the current point. The device **8** therefore computes: $0.25.\hat{Y}_{ij}+0.75.\hat{Y}_{ij}=\hat{Y}_{ij}$.

In the case of the points of the first line, spatial filtering is therefore reduced to horizontal filtering. During the time required to process the other lines of each field, the means **9** deliver a signal for controlling the multiplexer **35** in such a manner as to ensure that this latter connects the output of the register **34** to the input of the device **33**. The device **8** then computes a filtered value in accordance with the recurrence formula:

$$\hat{Y}_{ij}=0.25.\hat{Y}_{ij}+0.75.\hat{Y}_{i-1,j}.$$

This filtered value $\overline{Y}_{ij}$ constitutes the mean value M_v of luminance in the vicinity of the point considered.

The means **5** for multiplying the value Y_{ij} of luminance of the point being processed by a variable coefficient which is proportional to the ratio M_v/M_g are constituted by a read-only memory (ROM) **14** and three digital multipliers **10**, **11** and **12**. An address input of the read-only memory **14** is connected to the second input of the means **5** for receiving the mean value M_g. An output of the memory **14** delivers a value $1/M_g$ to a first input of the multiplier **10**. A second input of the multiplier **10** is connected to the first input of the means **5** for receiving a non-filtered luminance value Y_{ij}. An output of the multiplier **10** is connected to a first input of the multiplier **11** in order to deliver a value $Y_{ij}.1/M_g$ to this latter.

A second input of the multiplier **11** is connected to the third input of the means **5** for receiving the value M_v of the local mean. An output of the multiplier **11** is connected to a first input of the multiplier **12** in order to deliver a value $Y_{ij}.M_v/M_g$ to said first input. A second input of the multiplier **12** receives a constant value K which can be adjusted by an operator in order to adjust the contrast at will. An output of the multiplier **12** constitutes the output of the means **5** and is connected to the output terminal **13** of the contrast amplifier. Said output delivers a value $Y_{ij}.M_v/M_g.K$.

The coefficient M_v/M_g can be lower than or higher than 1 according to whether the vicinity of the point considered has a mean luminance value which is lower than or higher than the mean luminance value of the preceding field. Depending on which of these two cases applies, the effect of the contrast amplifier is therefore to reduce the luminance of the current point in order to bring it close to the value of black or respectively to increase said luminance in order to bring it close to the value of pure white.

Granular noise is represented by isolated grey points which are particularly visible in dark areas. As a result of bidirectional spatial filtering, the luminance values affected by noise produce very little disturbance in computation of the local mean value M_v and are consequently treated in the same manner as the values of the adjacent points. In the case of areas which are darker than the general mean value, these areas have an even darker appearance after processing including the points affected by noise. This is the contrary to what may be observed in the conventional method which consists simply in producing a uniform increase in gain in respect of all luminance values and therefore has the effect of increasing the luminosity of isolated grey points in the dark areas.

The invention is not limited to the example of construction described in the foregoing and many alternative embodiments are within the competence of those versed in the art. It is possible in particular to change the order of the horizontal filtering device **7** and of the vertical filtering device **8** or to change the order of the two multiplication operations which are necessary for computing $Y_{ij}.M_v/M_g.K$.

It is also possible to replace the means **6** by any other known device for computing a mean value of luminance in the vicinity of the point being processed.

The invention can be applied in particular to television pictures having low levels of illumination.

What is claimed is:

1. A contrast amplifier for video images, each video image being represented by the sequence of digital values of luminance of a points of said image, comprising:

first means for estimating a mean value M_g of luminance of all the points of each image in succession;

second means for computing a local mean value M_v of luminance in the vicinity of a point being processed;

means for multiplying the value of luminance of the point being processed by a variable coefficient which is proportional to the ratio M_v/M_g, the value thus obtained being such as to constitute the value of luminance of a point having enhanced contrast.

2. A contrast amplifier according to claim **1**, wherein the means for computing a local mean value M_v of luminance in the vicinity of a point comprise:

first recursive filtering means for receiving the sequence of luminance values X in order to compute a filtered value $\hat{X}$ in accordance with the formula:

$$\hat{X}=\alpha_1.X+(1-\alpha_1).A \text{ with } 0\leqq\alpha_1\leqq 1$$

where A is a filtered luminance value which has previously been computed by the first means in respect of one of the two points which are immediately adjacent to the point considered, namely the point which precedes said point on the same line and the corresponding point on the line scanned immediately beforehand;

and second recursive filtering means in series with the first means for computing the value M_v in accordance with the formula:

$$M_v=\alpha_2.\hat{X}+(1-\alpha_2).B \text{ with } 0\leqq\alpha_2\leqq 1$$

where B is a filtered luminance value which has previously been computed by the first and second means in respect of the other of the two points immediately adjacent to the point considered.

3. A contrast amplifier according to claim 2, wherein

the means for estimating a mean value M_g of luminance of all the points of each image comprise:

a first device for computing a mean value of luminance of the points on each line of an image,

a second device for computing a mean value of all the mean values computed by the first device in the case of each image.

* * * * *

Exhibit E

United States Patent [19]

Chen et al.

[11] **Patent Number: 4,789,933**

[45] **Date of Patent: Dec. 6, 1988**

[54] **FRACTAL MODEL BASED IMAGE PROCESSING**

[75] Inventors: **Victor C. Chen**, Richmond Hts., Ohio; **Mike M. Tesic**, Los Altos, Calif.

[73] Assignee: **Picker International, Inc.**, Highland Hts., Ohio

[21] Appl. No.: **19,568**

[22] Filed: **Feb. 27, 1987**

[51] **Int. Cl.**[4] **G06F 15/42**

[52] **U.S. Cl.** **364/413.13**; 382/47; 382/54

[58] **Field of Search** 382/27, 47, 54; 358/166, 167; 364/414

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,386,528	6/1983	Engle	358/112 X
4,633,503	12/1986	Hinman	382/47
4,665,551	5/1987	Sternberg	382/27 X
4,694,407	9/1987	Ogden	364/728 X
4,703,353	10/1987	David	382/47 X

OTHER PUBLICATIONS

"Analysis and Interpolation of Angiographic Images by Use of Fractals" by T. Lundahl, et al., IEEE, 1985, pp. 355–358.

"Digital Image Enchancement: A Survey" by Wang, et al. Computer Vision, Graphics, and Image Processing, 24, 363–381·(1983).

"Digital Image Processing by Use of Local Statistics" by Jong–Sen Lee, Naval Research Laboratory, pp. 55–61.

"Comparison of Interpolating Methods for Image Resampling" by Parker, et al., IEEE Transactions on Medical Imaging, vol. M1–2, No. 1, Mar. 1983, pp. 31–39.

"Nonstationary Statistical Image Models (and Their Application to Image Data Compression" by Hunt, 1980.

"Fractal Based Description of Natural Scenes" by Alex Pentland, IEEE 1983, pp. 201–209.

"A Note on Using the Fractal Dimension for Segmentation" by Medioni, et al, IEEE 1984, pp. 25–30.

"Computer Rendering of Stochastic Models" by Fournier, et al. Communications of the ACM, Jun. 1982, vol. 25, No. 6, pp. 371–384.

Primary Examiner—Jerry Smith

Assistant Examiner—Steven G. Kibby

Attorney, Agent, or Firm—Fay, Sharpe, Beall Fagan, Minnich & McKee

[57] **ABSTRACT**

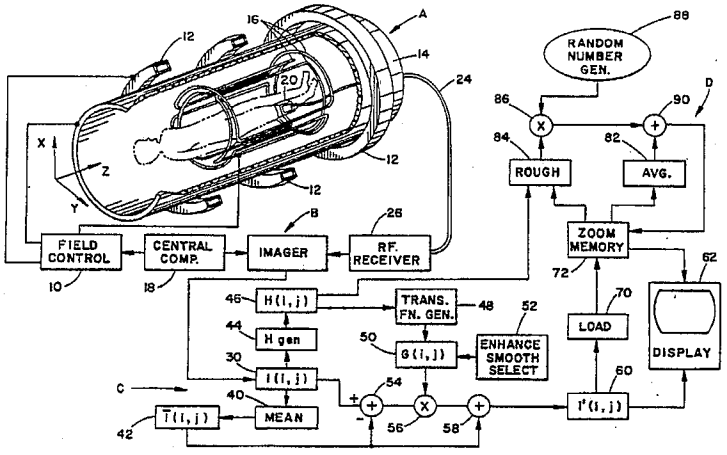
A medical diagnostic apparatus (A) generates medical diagnostic data which is reconstructed by an imager (B) into an electronic image representation. The electronic image representation includes an array of digital pixel values which represent a gray scale intensity of a man-readable image displayed on a video monitor (62). An image improving circuit (C) replaces each pixel value I(i,j) with an improved pixel value I'(i,j) defined as follows:

$$I'(i,j)=G(i,j)[I(i,j)-\bar{I}(i,j)]+\bar{I}(i,j),$$

where G(i,j) is a transfer function uniquely defined for each pixel and $\bar{I}$ is the mean of pixel values of neighboring pixels. The transfer function is based on a self-similarity value which is derived by comparing (i) a variation between the pixel value I(i,j) and pixel values in a first surrounding ring with (ii) a variation between the pixel value I(i,j) and pixel values in a second, larger surrounding ring. A zoom circuit (D) enlarges a selected portion of an improved image. Empty or unfilled pixel values I(inter) are interpolated from a combination of an average of neighboring pixel values I(avg) and a fractal value. The fractal value is a combination of a random number and a weighting factor R determined in accordance with the average self-similarity value of the neighboring filled pixels, i.e.:

$$I(\text{inter})=I(\text{avg})+R\cdot(\text{random}\#).$$

20 Claims, 2 Drawing Sheets



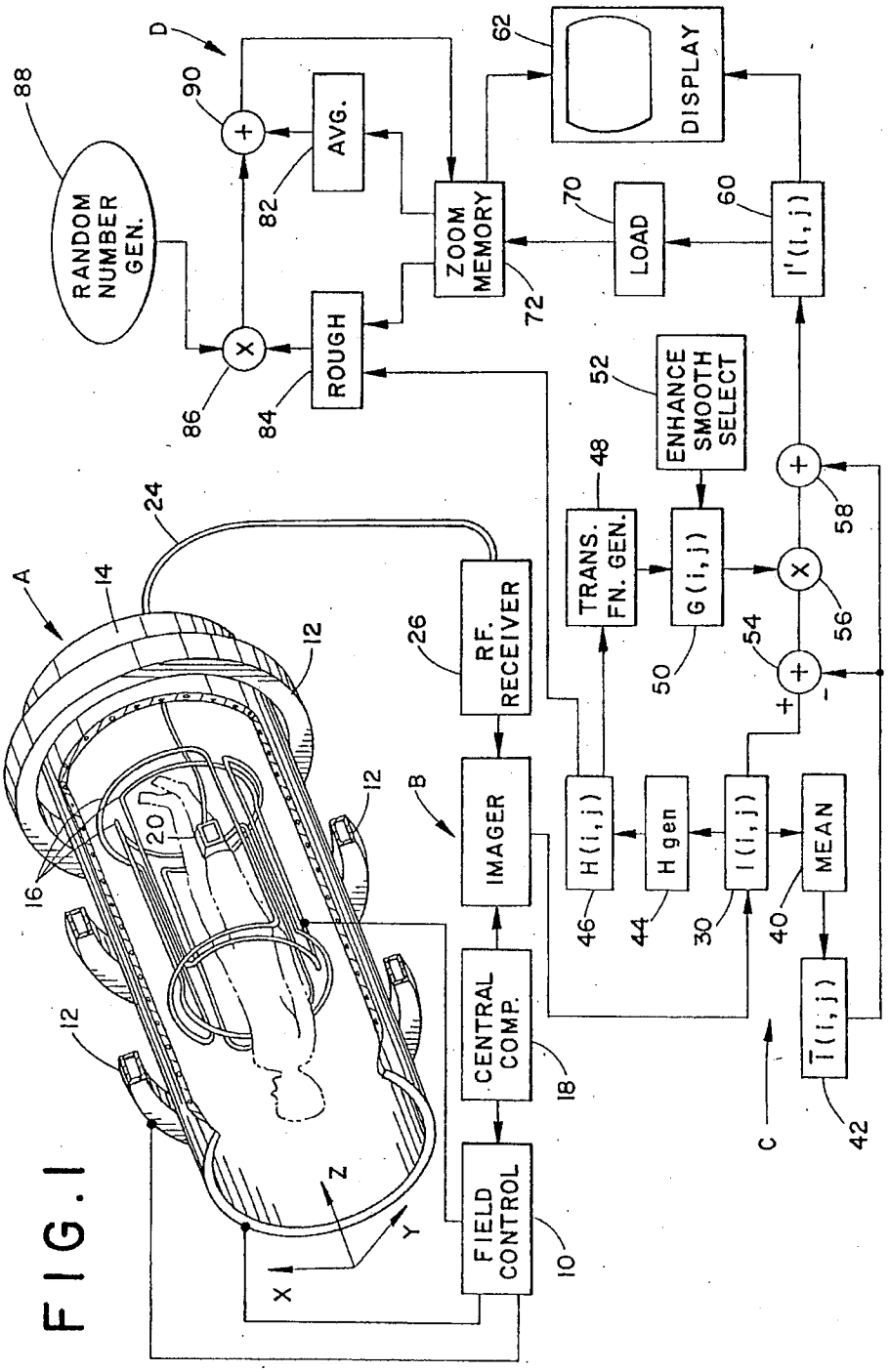

RANDOM NUMBER GEN.
88
D
90
AVG.
82
ROUGH
84
86
ZOOM MEMORY
72
LOAD
70
DISPLAY
62
I'(i,j)
60
ENHANCE SMOOTH SELECT
52
TRANS. FN. GEN.
48
G(i,j)
50
58
56
54
24
A
14
12
26
RF. RECEIVER
B
IMAGER
H(i,j)
46
H gen
44
I(i,j)
30
MEAN
40
Ī(i,j)
42
C
CENTRAL COMP.
18
FIELD CONTROL
10
16
20
Z
Y
X

FIG. 1

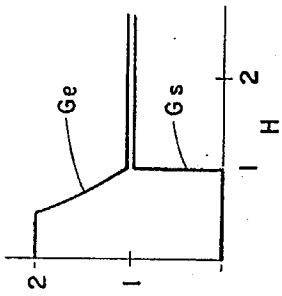


FIG. 3

FIG. 4

FIG. 2

FRACTAL MODEL BASED IMAGE PROCESSING

BACKGROUND OF THE INVENTION

The present invention relates to the art of image processing. It finds particular application in conjunction with image enhancement, image smoothing, image zooming and other image improvement techniques for magnetic resonance images and will be described with particular reference thereto. It is to be appreciated, however, that the present invention is also applicable to enhancing, improving and enlarging digital x-ray images, computed tomographic images, nuclear camera images, positron emission scanners, and the like.

Medical diagnostic images have commonly been subject to image degradation from noise, system imperfections, and the like. Various image processing techniques have been utilized to remove the effects of the noise and to highlight some specified features. See for example "Digital Image Enhancement: A Survey" Wang, et al., Computer Vision, Graphics, and Image Processing, Vol. 24, pages 363–381 (1983). In one technique, each pixel was adjusted in accordance with the mean of surrounding pixels and the variation or difference between the pixel value and the local mean (the average of the surrounding pixels). The enhanced pixel value g'(i,j) was a weighted average of the local mean and the variation:

$$g'(i,j)=\bar{g}(i,j)+k[g(i,j)-\bar{g}(i,j)] \quad (1),$$

where $\bar{g}(i,j)$ was the local mean, $g(i,j)-\bar{g}(i,j)$ was the variation, and k was a constant that weighted the relative contributions therebetween. It is to be appreciated that when k was set larger than 1, the variation, hence the fine details were magnified. As k was set smaller, the image was smoothed or blurred as if acted upon by a low-pass filter. At the extreme at which k was set equal to zero, each pixel value was replaced by the local mean of the neighboring pixel values.

One of the drawbacks in this technique resided in selecting an appropriate value for the weighting factor k. The smaller k was set, the more the image was blurred and the more difficult it became to extract accurate diagnostic information. As k was set larger, edges and fine details, including noise, became enhanced. Frequently, in a medical image, the selected weighting factor k was too large for some regions and too small for other regions.

"Digital Image Enhancement and Noise Filtering by Use of Local Statistics" by J. S. Lee, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 2, pages 165–168 (1980), recognized that different weighting factor k could be selected for each pixel to be enhanced. Specifically, Lee suggested setting the k for each pixel equal to the square root of the ratio of a preselected desirable local variance to the actual local variance of the selected pixel. Although Lee's weighting factor achieved better resultant images than the constant weighting factor, there was still room for improvement.

Another problem with medical diagnostic images resided in the blurring of enlarged or zoomed images. Typically, the diagnostic image included a fixed number of pixel values, e.g. a 256×256 pixel value matrix or array. When the image or a pattern thereof was enlarged each pixel could be displayed as a larger rectangle or additional intervening pixel values must be generated. For example, when a 256×256 array was enlarged to a 512×512 array, no data existed for alternate lines and for alternate columns of the 512×512 matrix. Commonly, the missing matrix values were interpolated by linear averaging the nearest neighboring pixel values. However, averaging adjacent pixels tended to blur the resultant enlarged image.

Another interpolation technique for zooming is described in "Analysis and Interpretation of Angiographic Images by Use of Fractals", T. Lundahl, et al., IEEE Computer in Cardiology, page 355–358 (1985). In Lundahl's technique, a global or image wide fractal model was utilized to derive the missing, intervening pixel values of an enlarged digital angiographic image. Lundahl, et al. first calculated a global fractal dimension for the entire image which described the average roughness or smoothness of the intensity surface of the entire image. The interpolated pixel values were based on the average of neighboring pixel values plus or minus a function of global fractal dimension. The plus or minus sign was chosen at random. One of the drawbacks in their technique is that the interpolated value was always different from the neighboring average by a function of the global fractal dimension when in fact, in the real world there is a certain chance for the interpolated value to be equal to the average of the neighboring values. While the addition or subtraction of a function of the global fractal dimension value made the image more realistic and easier to view, the interpolated pixel values were inaccurate and could cause an erroneous diagnosis. Particularly in low pixel value regions of an image with a large global fractal dimension, fictious image details could be generated during enlargement. This was caused by their use of the global fractal dimension in doing local interpolation. This is the other drawback of their technique.

The present invention is based on a new image fractal model, which describes images much more like real-world images. By the use of the fractal model, the processed medical image looks much more natural than the one by the use of traditional techniques.

The fractal model is based on the theory of fractal brownian motions developed by B. Mandelbrot. This theory provided a useful model for description of real-world surface. See 'Fractal: Form, Chance, and Dimension' and 'The Fractal Geometry of Nature', B. Mandrelbrot, W. H. Freeman and Company (1977) and (1982) respectively.

SUMMARY OF THE INVENTION

In accordance with one aspect of the present invention, a method of diagnostic imaging is provided. An electronic image representation which includes an array of pixel values is generated. A self-similarity value is generated corresponding to each pixel value. Each self-similarity value varies in accordance with the variations between the corresponding pixel value and the surrounding contiguous pixel values. Each pixel value of the image representation is replaced by a weighted combination of the pixel value to be replaced and the average of the surrounding pixel values. The combination is weighted in accordance with the corresponding self-similarlity value such that for enhancement the more similar the pixel value is to its surrounding contiguous pixel values, the more heavily the pixel value is weighted, i.e. the less enhancement is taken, and the less similar the pixel value is to its surrounding, contiguous

pixel values, the more heavily the variation value is weighted. However, for smoothing, the more similar the pixel value is to its surrounding contiguous pixel values, the more heavily the average of the surrounding pixel values is weighted and the less similar the pixel value is to its surrounding pixel values, the more heavily the pixel value is weighted.

More specific to the preferred embodiment, the self-similarity value is selected in accordance with a ratio of (i) a first averaging difference between the given pixel value and the pixel values in a first surrounding ring and (ii) a second averaging difference between the given pixel value and the pixel values in a second surrounding ring. In this manner, the self-similarity factor varies in accordance with the ratio of the averaging difference between the given pixel value and two concentric rings therearound.

In accordance with another aspect of the present invention, a method of diagnostic imaging is provided. An electronic image representation is defined by an array of pixel values, a fraction of the pixel values are generated from medical diagnostic data and the remaining pixel values are determined by interpolating the diagnostic data based pixel values. For each of a plurality of regions of the image, a roughness factor which varies with the degree of variation among pixel values within the region is determined. The roughness is largest in areas with large vatiations and smallest in homogeneous regions. Each interpolated pixel value is set equal to an average of surrounding diagnostic data based pixel values plus the product of the roughness factor and a Gaussian random number. In this manner, the interpolated pixel values are substantially equal to the average of neighboring diagnostic data based pixel values in substantially homogeneous regions. In regions with widely varying pixel values, the interpolated pixel values may vary more significantly from the neighboring value average.

In accordance with another aspect of the present invention, an apparatus is provided for processing diagnostic images. An image data memory means stores a plurality of pixel values based on diagnostic data. A self-similarity value calculating means calculates a self-similarity value corresponding to each pixel value stored in the memory means. The self-similarity values are determined in accordance with a ratio of the differences between the corresponding pixel value and the pixel values in at least two surrounding rings of pixel values. A mean value determining means determines for each pixel value a corresponding mean value based on the average of neighboring pixel values. A filtering means combines each pixel value with the corresponding mean value with the relative weight of the pixel value and the corresponding mean value being weighted in accordance with the corresponding self-similarity value.

In accordance with another aspect of the present invention, an apparatus for processing diagnostic images is provided. A diagnostic imaging means generates a plurality of diagnostic data based pixel values which are stored at a predetermined fraction of the pixel of an image memory. The unfilled image memory pixels are denoted as interpolated pixels. An averaging means averages diagnostic data based pixel values neighboring each interpolated pixel. A roughness factor determining means determines a roughness factor in the neighborhood of each interpolated pixel, which roughness factor varies in accordance with the variation of diagnostic data based pixel values in the region. A random number generator generates Gaussian random numbers. A combining means combines the random number generated by the random number generator with the roughness factor to produce a fractal value and combines the fractal value with the average from the averaging means to generate an interpolated pixel value. An interpolating means returns each interpolated pixel value to the image memory for storage at the corresponding interpolated pixel.

One advantage of the present invention is that it generates medical diagnostic images of improved viewability.

Another advantage of the present invention is that it provides an improved pixel selective filtering function for automatically smoothing noise an enhancing resultant images.

Another advantage of the present invention is that it enlarges images while retaining the same apparent sharpness to the viewer's eye.

Still further advantages of the present invention will become apparent to those of ordinary skill in the art upon reading and understanding the following detailed description of the preferred embodiment.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention may take form in various steps and arrangements of steps and in various components and arrangements of components. The drawings are for purposes of illustrating a preferred embodiment and are not to be construed as limiting the invention.

FIG. **1** is a diagrammatic illustration of a medical diagnostic imaging system in accordance with the present invention;

FIG. **2** diagrammatically illustrates calculation of a self-similarity factor in a 5×5 pixel region about a pixel (i,j);

FIG. **3** is a diagrammatic illustration of a preferred transfer function G(i,j); and,

FIG. **4** is a diagrammatic illustration of an interpolation technique around the pixels of the zoom memory.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference to FIG. **1**, a medical diagnostic apparatus A generates medical diagnostic data which is reconstructed by an imager B into an electronic image representation. An image filter and enhancement circuit C operates on the electronic image representation to improve the image quality and viewability thereof. A zoom circuit D selectively enlarges the resultant image representation or regions thereof.

Although a magnetic resonance imager is illustrated, the medical diagnostic apparatus A may be a computerized tomographic scanner, a digital x-ray apparatus, a positron emission scanner, a nuclear camera, or other diagnostic apparatus which generates data that is able to be reconstructed into an image representative of a region of an examined patient or subject. The illustrated magnetic resonance imager includes a field control means **10** which controls a main, homogeneous polarizing magnetic field through an image region generated by electromagnets **12**. The field control means **10** also controls gradient magnetic fields created across the image region by gradient field coils **14** to provide spatial encoding, phase encoding, and slice select gradients. The field control means **10** further generates radio frequency electromagnetic excitation signals which are

applied to excitation coils **16** to excite resonance of dipoles in the image region. A central computer **18** controls the relative timing and strengths of the gradient and radio frequency electromagnetic fields.

A receiving coil **20** is surface mounted on the patient to receive magnetic resonance signals generated by resonating dipoles in the image region. A flexible cable **24** conveys the magnetic resonance signals to radio frequency receiver **26**. The radio frequency receiver **26** supplies magnetic resonance medical diagnostic data to the imager B.

The imager B under control of the central computer **18** reconstructs the medical diagnostic data into an electronic image representation. More specifically, the imager B reconstructs digital pixel values, each pixel value corresponding to a preselected subregion or pixel of the image region. The algorithm implemented by the imager is selected in accordance with the medical diagnostic apparatus selected. Conventionally, the pixels or subregions are arranged in a rectangular array (i,j) with the corresponding pixel values I(i,j) being arranged in a like rectangular array. The electronic image representation, particularly the rectangular array of pixel values I(i,j), is stored in a first image memory **30**.

The image filtering and enhancing circuit C includes a pixel value averaging means **40** which generates a matrix of averaged or mean pixel values for storage in a mean value memory **42**. Each mean pixel value I(i,j) is set equal to the average of the pixel values in the neighborhood of a corresponding pixel value of the first image memory **30**. Preferably, the pixel mean value means **40** determines each pixel mean value as follows:

$$I(i,j) = \frac{1}{(2l+1)^2} \sum_{x=0}^{l} \sum_{y=0}^{l} I(i \pm x, j \pm y), \tag{2}$$

where l is indicative of the size of the neighborhood over which the mean value is averaged. For example, as illustrated in FIG. **2**, for l=2, a 5×5 neighborhood is defined in which the mean value represents the average of the 25 pixel values in the neighborhood. Optionally, other averaging algorithms may be utilized. For example, a 3×3 neighborhood may be selected. As yet another option, the average may include the average of the surrounding pixel values in the neighborhood but not the pixel value itself, e.g. the 24 surrounding pixel values of a 5×5 neighborhood. The surrounding pixel values may be weighted inversely with radial distance, or the like.

A self-similarity parameter generator **44** derives a self-similarity value H(i,j) corresponding to each pixel value for storage in a self-similarity value memory **46**. In Euclidean geometry, th Euclidean dimension, E, is an integer. For example, E=2 for lines and E=3 for surfaces. In topological geometry, the topological dimension, Dt, is also an integer, but Dt=1 for lines and Dt=2 for surfaces. In fractal geometry, the fractal dimension, D, need not be an integer, although it is commonly a real number. The fractal dimension is equal to or greater than the topological dimension and less than corresponding Euclidian dimension. For example, $1 \leqq D \leqq 2$ for lines and $2 \leqq D < 3$ for surfaces.

The fractal dimension may be viewed as a measurement of the degree of irregularity or roughness. As the fractal dimension becomes larger and approaches the Euclidean dimension, the surface becomes more rough; as the fractal dimension becomes smaller and approaches the topological dimension, the surface becomes smoother.

The self-similarity parameter H is equal to the difference in the Euclidean and fractal dimensions, i.e.

$$H = E - D \tag{3}.$$

Thus, the smoothest surface is defined when H=1 and the surfaces become rougher as H becomes less than 1.

In the present invention, the self-similarity parameter for each pixel, H(i,j) is determined by monitoring the neighborhood or environment around the corresponding pixel. The greater the variation between a given pixel value and the surrounding pixel values, the smaller the self-similarity parameter. The more uniform the neighborhood, the closer the self-similarity parameter approaches one.

With reference to FIG. **2**, the self-similarity parameter is derived using a convolution type function. Specifically, for a given pixel, (i,j), the averaging difference between the given pixel value I(i,j) and pixel values in a first surrounding ring is compared with the averaging difference between the given pixel value and pixel values in a second surrounding ring. For rings of radius m and n, the definition of a scaler Brownian function provides:

$$(n/m)^H = \frac{k(n;i,j)}{k(m;i,j)}, \tag{4}$$

$$k(n;i,j) = \left(\sum_{(n-1)<r \leqq n} |I(i,j) - I(k,l)| \right) / N, \tag{5}$$

where

N=Number of pixels whose centers fall in the ring between n and n−1.

and

$$r^2 = (i-k)^2 + (j-l)^2 \tag{6}.$$

An expression of H can be obtained by taking the base-10 logarithm of both sides of Equation (4):

$$H \log(n/m) = \log\left(\frac{k(n;i,j)}{k(m;i,j)} \right) \tag{7}$$

or

$$H(i,j) = \frac{\log k(n;i,j) - \log k(m;i,j)}{\log(n) - \log(m)} \tag{8}$$

Thus, a self-similarity parameter or value H is determined in accordance with a ratio of two differences. The two differences are (1) the difference between the logarithm of the function k at the nth ring and the logarithm of the function k at the mth ring and (2) the difference between the logarithm of n and the logarithm of m.

In the embodiment illustrated in FIG. **2**, n=2 and m=1. For pixel (i,j), illustrated in FIG. **2**, k(2;i,j) is the average of the differences between the pixel value I(i,j) and the pixel values corresponding to each of pixels (i,j+2), (i−1,j+1), (i−2,j), (i−1,j−1), (i,j−2), (i+1,j−1), (i+2,j), and (i+1,j+1). Similarly, with the radius of the inner ring set equal to one, k(1;i,j) includes the average of the differences between the intensity at

pixel (i,j) and the intensities corresponding to each of pixels (i,j+1), (i−1,j), (i,j−1), and (i+1,j). That is, k(1;i,j) is the average of the differences between the pixel intensity or pixel value of pixel (i,j) and each of the pixels whose centers line on or within the radius of one interpixel spacing, r=1, and k(2;i,j) includes the difference between the value of pixel (i,j) and the pixel value corresponding to each pixel whose center lies outside of the r=1 ring and on or inside the r=2 ring.

It is to be appreciated that although the two rings are illustrated in the preferred embodiment as being continuous, a gap may be defined between the two rings. For example, the first ring might include those pixels whose centers fall between ¾ and 1¼ of the interpixel spacing and the second ring might include those pixels whose centers fall in the ring between 1¾ and 2¼ times the interpixel spacing. As another example, the pixel value of each pixel whose rectangular area falls even partially within a given ring may be determined and the difference may be weighted in accordance with the percentage of the pixel area which falls within the given ring. Optionally, additional rings may be defined concentrically around the first and second rings. Preferably, the effect of additional outer rings is reduced inversely with the larger radius of the outer rings. In this manner, the self-similarity value for each pixel is derived from a convolution type function.

With continuing reference to FIG. **1** and further reference to FIG. **3**, a transfer function means **48** derives a transfer function for each pixel G(i,j) which may be stored in a transfer function memory means **50**. In the preferred embodiment, the transfer function may either (i) enhance the image, i.e. accentuate edge effects or (ii) smooth the image, i.e. reduce noise effects for a more uniform image. For enhancement, an enhancement transfer function $G_e(i,j)$ is utilized. In the preferred embodiment, the enhancement transfer function is selected in accordance with:

$$G_e(i,j) = \begin{cases} 2 & \text{for } H(i,j) < 0.5 \\ 1/H(i,j) & \text{for } 0.5 \leqq h(i,j) \leqq 1.0 \\ 1 & \text{for } H(i,j) > 1.0 \end{cases} \quad (9)$$

For smoothing, a smoothing transfer function $G_s(i,j)$ is selected for each pixel in accordance with:

$$G_s(i,j) = \begin{cases} 1 & \text{for } H(i,j) > 1.0 \\ 0 & \text{for } H(i,j) \leqq 1.0 \end{cases} \quad (10)$$

An operator control **52** enables the operator to select between the described enhancement and smoothing transfer functions or other preselected transfer functions.

It is to be appreciated that the image memory **30**, the mean value memory **42**, the self-similarity value memory **46**, and the transfer function memory **50** are for purposes of illustration. The values described as stored in these memories may be stored as described, calculated in real time, or a combination of both.

In the preferred embodiment, the image smoothing-/enhancement circuit C derives an improved image value for each pixel, I'(i,j), which is related to a combination of the pixel value, I(i,j) and the corresponding mean pixel neighborhood value $\bar{I}(i,j)$ weighted in accordance with the transfer function G(i,j), specifically:

$$I'(i,j) = G(i,j)[I(i,j) - \bar{I}(i,j)] + \bar{I}(i,j) \quad (11).$$

A subtraction means **54** subtractively combines pixel by pixel the pixel intensity I(i,j) and the mean neighborhood value $\bar{I}(i,j)$. A multiplying means **56** multiplies the difference between the pixel and mean neighborhood values by the corresponding transfer function value G(i,j). An adding means **58** additively combines the mean neighborhood value and the transfer function product pixel by pixel. Each resultant enhanced/filtered pixel value I'(i,j) is stored at the corresponding pixel of a second or improved image memory **60**. The improved image may be displayed on a video monitor **62** or other display means. Optionally, the image may be stored on magnetic tape or disk subject to further processing, or the like.

With continuing reference to FIG. **1** and further reference to FIG. **4**, if a portion of the resultant image is to be enlarged for closer review, the zoom circuit D is activated. In the enlarged or zoomed image, there are more pixels than there are pixel values in the original image from the improved image memory **60**. A subregion loading means **70** loads the pixel values from the region to be enlarged in the original image into corresponding pixels which are distributed uniformly over a zoom memory **72**. If the region of the original image is to be doubled in width and height, then there are empty pixels, note pixels **74** and **76**, between each filled pixel of each column. Analogously, between each row of filled pixels, the zoom memory **72** has an empty row of pixels, note rows **78** and **80**. It is to be appreciated that if the region of interest is to be more enlarged, there will be more intervening unfilled pixels. If the region of interest is to be less enlarged, there will be fewer unfilled pixels. The zoom means D derives pixel values for each empty pixel in the zoom memory **72** from the pixel values loaded from the original image **60** by the loading means **70**. More specifically, the zoom circuit D fills each empty pixel with a value that is equal to the sum of an average of neighboring pixel values and a fractal value F.

An averaging means **82** averages the pixel values of the filled pixels which neighbor or are contiguous to each empty pixel. For example, empty pixel **74** is filled with the average of the neighboring pixel values (i,j) and (i+1,j), i.e. $[I(i,j)+I(i+1,j)]/2$. Optionally, the average may be based on other average or mean value algorithms.

An image may be thought of as an intensity surface, where each pixel intensity value represents height (z-coordinate) above the corresponding pixel location on x-y plane.

The fractal value F is a function of the roughness or variation in neighboring values and random numbers. A roughness means **84** calculates a roughness factor R for each empty pixel based on the self-similarity factors of neighboring filled pixels. Preferably, the roughness factor is set equal to a power of the average self-similarity factor of the adjoining filled pixels. For pixel **74**, the roughness factor is:

$$R = h.2^{-[H(i,j)+H(i+1,j)]/2} \quad (12),$$

where h is a constant selected in accordance with system parameters, the video monitor, and the desired

characteristics of the zoomed image. Again, the exponent may be based on the nearest neighbors and more distant neighboring pixels. A multiplying means **86** multiplies the roughness factor R for each interpolated pixel by a random number generated by random number generator **88** to generate the fractal value F. An adding means **90** combines the fractal value and the average value to produce the interpolated pixel value, i.e.:

$$I(\text{inter})=[I(i,j)+I(i+1,j)]/2+R\cdot(\text{random \#}) \quad (13).$$

The interpolated value for pixel **78** may be calculated analogously based on filled pixel values I(i,j) and I(i,j+1). The interpolated value for a pixel **92** at the intersection of empty rows and columns may be determined from the intensities and self-similarity factors from the four nearest neighbor pixels (i,j), (i+1,j), (i,j+1), and (i+1, j+1). Alternately, the pixel value **92** may be interpolated analogously from the four nearest interpolated pixel values or from a combination of the four nearest interpolated pixel values and the four nearest filled pixel values.

The invention has been described with reference to the preferred embodiments. Obviously, modifications and alterations will occur to others upon reading and understanding the preceding detailed description. It is intended that the invention be construed as including all such modifications and alterations insofar as they come within the scope of the appended claims or the equivalents thereof.

Having thus described the preferred embodiment, the invention is now claimed to be:

1. An apparatus for generating medical diagnostic image representations, the apparatus comprising:

a medical diagnostic apparatus for generating diagnostic data indicative of a preselected image region of a patient;

an imager for generating an electronic image representation from the diagnostic data, the electronic image representation including a pixel value corresponding to each pixel of a pixel array, each pixel value is indicative of an image property of a corresponding subregion of the image region of the patient;

a self-similarity value generating means for generating a self-similarity value corresponding to each pixel, the self-similarity value generating means being operatively connected with the imager to receive pixel values therefrom, each self-similarity value varying in accordance with a ratio of (i) a difference between the corresponding pixel value and a first set of pixel values contiguous to and surrounding the corresponding pixel value and (ii) the corresponding pixel value and a second set of pixel values contiguous to and surrounding the first set of pixel values;

an image improvement means for replacing each pixel value by a combination of the replaced pixel value and an average of surrounding pixel values, the combination being weighted in accordance with the corresponding self-similarity value, the image improvement means being operatively connected with the imager and the self-similarity value generating means.

2. An apparatus for generating medical diagnostic image representations, the apparatus comprising:

a medical diagnostic apparatus for generating medical diagnostic signals indicative of an image region of a patient;

an imager for transforming the medical diagnostic signals into electronic image representations, each electronic image representation including a pixel value corresponding to each pixel of a pixel array;

a loading means for loading each pixel value into uniformly distributed, spaced pixels of a zoom memory means such that the pixel values are stored in uniformly distributed filled pixels and empty pixels are disposed therebetween;

an interpolating means for generating a pixel value for each empty pixel, the interpolating means including:

an averaging means for generating an average of neighboring pixel values;

a random number generator for generating a random number;

a self-similarity value generating means for generating a self-similarity value corresponding to each pixel, the self-similarity value generating means being operatively connected with the imager to receive pixel values therefrom, each self-similarity value varying in accordance with a ratio of differences between the corresponding pixel value and pixel values in at least two rings contiguous to the corresponding pixel value;

a weighting means for weighting the random number in accordance with the self-similarity value, the weighting means being operatively connected with the self-similarity value generating means and the random number generator; and,

combining means for combining the pixel value average from the averaging means and the weighted randon number from the weighting means, the combining means being operatively connected with the averaging means to receive the average value therefrom, with the weighting means for receiving the weighted random number therefrom, and with the zoom memory means for storing the combined average pixel value and weighted random number at the corresponding empty pixel thereof.

3. A method of medical diagnostic imaging comprising:

converting medical diagnostic data into an electronic image representation which includes an array of pixel values;

generating a self-similarity value corresponding to each pixel value of the electronic image representation, each self-similarity value being a dimensionless value which varies with a degree of irregularity among pixel values in a neighborhood of the corresponding pixel value;

providing an improved electronic image representation by replacing each pixel value by a combination of the replaced pixel value and an average of surrounding pixel values, the combination being weighted in accordance with the corresponding self-similarity value.

4. The method as set forth in claim **3** wherein the self-similarity value generating step includes:

determining a difference between the corresponding pixel value and each pixel value within a nearest neighbor first ring and averaging the first ring differences;

determining a difference between the corresponding pixel value and pixel values in a next nearest neighboring second ring and averaging the second ring differences.

5. The method as set forth in claim 4 wherein the self-similarity value generating step further includes determining a logarithm of a ratio of the first ring and second ring differences.

6. The method as set forth in claim 5 further including multiplying the logarithm by a constant selected in accordance with the diameter of the rings.

7. The method as set forth in claim 5 wherein the replacing step further includes determining a transfer function for each pixel value from the corresponding self-similarity value and wherein the combination is weighted by multiplying the combination by the transfer function.

8. The method as set forth in claim 7 wherein the transfer function is inversely proportional to the self-similarity value for at least a selected range of self-similarity values.

9. The method as set forth in claim 7 wherein the transfer function is set equal to a first constant when the self-similarity value exceeds a preselected value and the transfer function is set equal to a second preselected constant when the self-similarity value is equal to or less than the preselected value.

10. The method as set forth in claim 4 further including enlarging a selected portion of the electronic image representation, the enlarging step including distributing pixel values from the selected image representation portion substantially uniformly among available pixels of an enlarged image electronic representation such that a fraction of the enlarged image representation pixels are filled by the transferred pixel values and empty pixels are defined therebetween and interpolating neighboring filled pixel values to derive pixel values for the empty pixels.

11. The method as set forth in claim 10 wherein the interpolating step includes:

for each empty pixel, averaging pixel values from neighboring filled pixels; and,

adding to the averaged pixel valves a fractal value which varies in accordance with a random number and the self-similarity value.

12. The method as set forth in claim 11 wherein the fractal value varies in proportion to a constant raised to a power of an average of the self-similarity values corresponding to the neighboring filled pixels.

13. A method of improving an image comprising:

converting data into an electronic image representation which includes an array of pixel values;

determining a first difference between the corresponding pixel value and each pixel value within a first surrounding ring and averaging the first ring differences;

determining a difference between the corresponding pixel value and each pixel value in a second surrounding ring and averaging the second ring differences;

generating a self-similarity value corresponding to each pixel value of the electronic image representation in accordance with a difference between the first and second difference averages corresponding to the same pixel value;

providing an improved electronic image representation by replacing each pixel value by a combination of the replaced pixel value, pixel values of pixels surrounding the replaced pixel value, and the self-similarity value corresponding to the replaced pixel value.

14. A method of medical diagnostic imaging comprising:

converting medical diagnostic data into an electronic image representation which includes an array of pixels, each pixel having a pixel value;

generating a self-similarity value corresponding to each pixel;

enlarging a selected portion of the electronic image representation, the enlarging step including distributing pixel values of the selected image representation portion substantially uniformly among pixels of an enlarged image electronic representation such that a fraction of the enlarged image representation pixels are filled by the transferred pixel values and empty pixels are defined therebetween;

for each empty pixel, averaging pixel values from neighboring filled pixels;

for each empty pixel deriving a fractal value which varies in accordance with a random number and the self-similarity values of the neighboring filled pixels; and,

combining the fractal value and the neighboring pixel value average that corresponds to the same empty pixel.

15. The method as set forth in claim 14 wherein the fractal value varies in proportion to a constant raised to a power of an average of the self-similarity values corresponding to the neighboring filled pixels.

16. The method as set forth in claim 15 wherein each self-similarity value varies in accordance with a ratio of variations between the corresponding pixel value and pixel values in at least two surrounding rings.

17. The method as set forth in claim 15 wherein the self-similarity value generating step includes:

determining a difference between the corresponding pixel value and each pixel value within a first surrounding ring and averaging the first ring differences;

determining a difference between the corresponding pixel value and each pixel value in a second surrounding ring and averaging the second ring differences.

18. The method as set forth in claim 17 wherein the self-similarity value generating step further includes determining a logarithm corresponding to a ratio of the first ring and second ring differences.

19. The method of claim 18 including providing an improved electronic image representation by replacing each pixel value by a combination of the pixel value and an average of surrounding pixel values, each combination being weighted in accordance with the corresponding self-similarity value.

20. A method of enlarging a selected portion of an image representation comprising:

converting data into an electronic image representation which includes an array of pixels, each pixel having a pixel value;

generating a self-similarity value corresponding to each pixel from pixel values of adjacent pixels;

distributing pixel values of the selected image representation portion substantially uniformly among pixels of an enlarged image electronic representation such that a fraction of the enlarged image representation pixels are filled by the transferred

pixel values and empty pixels are defined therebetween;

for each empty pixel, averaging pixel values from neighboring filled pixels;

for each empty pixel, deriving a fractal value which varies in accordance with a random number and the self-similarity values of adjacent pixels; and,

combining the fractal value and the neighboring pixel value average that corresponds to the same empty pixel.

* * * * *

Exhibit F

Digital Image Processing

APPLIED MATHEMATICS AND COMPUTATION

A Series of Graduate Textbooks, Monographs, Reference Works

Series Editor: ROBERT KALABA, University of Southern California

No. 1 MELVIN R. SCOTT
Invariant Imbedding and its Applications to Ordinary Differential Equations: An Introduction, 1973

No. 2 JOHN CASTI and ROBERT KALABA
Imbedding Methods in Applied Mathematics, 1973

No. 3 DONALD GREENSPAN
Discrete Models, 1973

No. 4 HARRIET H. KAGIWADA
System Identification: Methods and Applications, 1974

No. 5 V.K. MURTHY
The General Point Process: Applications to Bioscience, Medical Research, and Engineering, 1974

No. 6 HARRIET H. KAGIWADA and ROBERT KALABA
Integral Equations Via Imbedding Methods, 1974

No. 7 JULIUS T. TOU and RAFAEL C. GONZALEZ
Pattern Recognition Principles, 1974

No. 8 HARRIET H. KAGIWADA, ROBERT KALABA, and SUEO UENO
Multiple Scattering Processes: Inverse and Direct, 1975

No. 9 DONALD A. PIERRE and MICHAEL J. LOWE
Mathematical Programming Via Augmented Lagrangians: An Introduction with Computer Programs, 1975

No. 10 J. SUSAN MILTON and CHRIS P. TSOKOS
Probability Theory with the Essential Analysis, 1976

No. 11 WOLFGANG EICHHORN
Functional Equations in Economics, 1978

No. 12 SIDNEY J. YAKOWITZ
Computational Probability and Simulation, 1977

No. 13 RAFAEL C. GONZALEZ and PAUL WINTZ
Digital Image Processing, 1977

Other Numbers in Preparation

Digital Image Processing

Rafael C. Gonzalez

Department of Electrical Engineering
University of Tennessee, Knoxville

Paul Wintz

School of Electrical Engineering
Purdue University, Lafayette, Indiana
and
Wintek Corporation, Lafayette, Indiana

1977
Addison-Wesley Publishing Company
Advanced Book Program
Reading, Massachusetts

London · Amsterdam · Don Mills, Ontario · Sydney · Tokyo

Library of Congress Cataloging in Publication Data

Gonzalez, Rafael C
Digital image processing.

(Applied mathematics and computation ; no. 13)
Includes bibliographies and index.
1. Image processing. I. Wintz, Paul A., joint author. II. Title.
TA1632.G66 621.38'0414 77-10317
ISBN: 0-201-02596-5
ISBN: 0-201-02597-3 pbk.

Reproduced by Addison-Wesley Publishing Company, Inc., Advanced Book Program, Reading, Massachusetts, from camera-ready copy prepared under the supervision of the authors.

American Mathematical Society (MOS) Subject Classification Scheme (1970): 68-00, 68A45, 62-04, 65C99, 42-00, 42A68, 42A76, 92-04, 93-00, 94-04, 94A05, 94A10.

APPENDIX A

IMAGE DISPLAY SUBROUTINE

The following FORTRAN subroutine outputs to a line printer a 32-level image. The maximum array size is a 64 X 64, which produces a full, standard computer output sheet. Larger images can be printed by subdividing them into 64 X 64 mosaics.

Usage

IA - A 64 X 64 image. This integer array is formed in the calling program where it is set up as follows:
COMMON IA (64, 64)
INTEGER*2 IA
The first subscript of IA refers to a *row* of the image and the second to a *column*. Care must be taken to keep this convention in mind when building the array.

The following subroutine displays IA on a line printer.

CALL DSP(NX,NY,LAW,IL,IH,NEG,LG)

The arguments are as follows:

NX - Number of rows of IA to be printed; maximum NX is 64.
NY - Number of columns of IA to be printed; maximum NY is 64. If NX = NY = 64 a full page is output.
LAW - Gray-level scale translation variable.
LAW = 1: linear scale,
LAW = 2: square-root scale,
LAW = 3: logarithmic scale,
LAW = 4: "absorption" scale.

IL - minimum gray level in IA, calculated in the calling program.
IH - maximum gray level in IA, calculated in the calling program.
NEG - a value equal to 1 gives the normal image; a value equal to 0 gives the negative of the image.
LG - logical unit number for the line printer.

The characters used in the program to obtain the thirty-two gray levels are shown in Fig. A.1. The characters in a column, when over-printed, produce the gray level indicated.

```
MMMMMMMHHHHHHXHXOZWMNOS=I*++=:-.-
WWWWWW###*++----      =   -    -
####0O+-
O0O
+

                                    Gray Levels
```

Figure A.1. Over-print characters used to obtain thirty-two gray levels. The 32nd character is a blank.

```
      SUBROUTINE DSP(NX,NY,LAW,IL,IH,NEG,LG)
C
C     **LINE PRINTER IMAGE OUTPUT SUBROUTINE**
C
C     ADAPTED BY B. A. FITTES, ELECTRICAL ENG.
C     DEPT., UNIVERSITY OF TENNESSEE, FROM
C     A PROGRAM WRITTEN BY J. L. BLANKENSHIP,
C     INSTRUMENTATION AND CONTROLS DIV., OAK
C     RIDGE NATIONAL LABORATORY, OAK RIDGE,
C     TN.  THE OVERPRINT METHOD USED IS FROM
C     "CONSIDERATIONS FOR EFFICIENT PICTURE
C     OUTPUT VIA LINE PRINTER," BY P. HENDERSON
C     AND S. TANIMOTO, REPORT NO. 153, 1974,
C     COMPUTER SCIENCE LAB., ELEC. ENG. DEPT.,
C     PRINCETON UNIVERSITY.
      COMMON IA(64,64)
      INTEGER*2 IA
      INTEGER*2 IB(64,64),LEV(32),BLANK(5)
```

```
      LOGICAL*1 LINE(128,5),GRAY(32,5)
C     SPECIFY GRAY-LEVEL CHARACTERS
      DATA GRAY /6*'M',5*'H','X','H','X','O',
     1 'Z','W','M','N','O','S','=','I','*','+',
     2 '+','=',':','-','.','-',' ',6*'W',3*'#',
     3 '*','+','+',4*'-',5*' ','=',' ',' ','-',
     4 3*' ','-',3*' ',4*'#','O','O','+','-',
     5 24*' ','O','O','O',29*' ','+',31*' '/
      GN=32.0
      FL=IL
      FH=IH
      IF ((FH-FL).GT.0.0) GO TO 100
      T=FL
      FL=FH
      FH=T
  100 RANGE=(FH-FL+1)/GN
      AA=(SQRT(FH)-SQRT(FL))/GN
      EE=(FH-FL)/ALOG(GN+1.0)
      T=AMAX1(FL,1.0)
      SS=-(1.0/GN)*ALOG(FH/T)
C
C
C     A VECTOR LEV IS COMPUTED NEXT.  THIS
C     VECTOR IS A SET OF BREAK POINTS USED TO
C     DETERMINE THE SCALED VALUE OF IA(I,J).
C     THE MATRIX IB IS THE RESULT OF SCALING
C     IA.  IB(I,J)=K IF IA(I,J) IS LESS THAN
C     LEV(K+1) BUT GREATER THAN OR EQUAL TO
C     LEV(K).
C
      DO 160 I=1,32
      GO TO (110,120,130,140),LAW
  110 FLEV=FL+(I-1)*RANGE+0.5
      GO TO 150
  120 FLEV=(SQRT(FL)+(I-1)*AA)**2+0.5
      GO TO 150
  130 FLEV=FL+EE*ALOG(FLOAT(I))+0.5
      GO TO 150
  140 FLEV=FH*EXP(SS*(GN-I))+0.5
  150 LEV(I)=FLEV
  160 CONTINUE
      IF (NX.GT.64) NX=64
      IF (NY.GT.64) NY=64
      DO 180 I=1,NX
      DO 180 J=1,NY
      KLT=1
      DO 170 K=1,32
```

```
      IF (IA(I,J).GE.LEV(K)) KLT=K
  170 CONTINUE
      IB(I,J)=KLT
  180 CONTINUE
C
C
C     ONCE IB HAS BEEN COMPUTED, THE PICTURE CAN
C     BE PRINTED.  EACH POINT IN THE PICTURE CAN
C     CONSIST OF UP TO FIVE CHARACTERS OVERPRINTED
C     ON ONE ANOTHER.  SINCE THERE ARE 32 POSSIBLE
C     GRAY LEVELS, THERE IS A 32X5 MATRIX, GRAY,
C     THAT CONTAINS ALL OF THE COMBINATIONS.  SINCE
C     EACH ELEMENT OF IB IS AN INTEGER BETWEEN
C     1 AND 32, IT CAN BE USED AS AN INDEX ON
C     GRAY TO OBTAIN THE CORRECT COMBINATION.
C     THE OUTPUT BUFFER, LINE, IS A 128X5 MATRIX.
C     EACH POINT IS OUTPUT TWICE HORIZONTALLY AND
C     ONCE VERTICALLY TO ATTEMPT TO COMPENSATE
C     FOR THE SPACING OF THE PRINTER.  HENCE, THE
C     FULL OUTPUT BUFFER REPRESENTS ONE ROW OF IB.
C     AS THE ROW IS GENERATED THERE IS A VECTOR,
C     BLANK, THAT INDICATES WHETHER OR NOT ANY
C     NON-BLANK CHARACTERS ARE PRESENT IN THE
C     BUFFER.  IF THERE ARE NOT ANY, THAT ROW IS
C     NOT PRINTED.  THIS SPEEDS UP THE PRINTING
C     PROCESS.
C
      WRITE(LG,1)
      IX=NX
      IY=2*NY
      DO 210 I=1,IX
      DO 190 J=1,5
      BLANK(J)=0
  190 CONTINUE
      DO 200 K=2,IY,2
      J=K/2
      NG=IB(I,J)
      IF (NEG.EQ.0) NG=33-NG
      DO 200 L=1,5
      LINE(K-1,L)=GRAY(NG,L)
      LINE(K,L)=GRAY(NG,L)
      IF (NG.NE.32) BLANK(L)=1
  200 CONTINUE
      WRITE (LG,2)
      DO 210 L=1,5
      IF (BLANK(L).EQ.0) GO TO 210
      WRITE (LG,3) (LINE(M,L), M=1,IY)
```

```
210 CONTINUE
    1 FORMAT (1H1)
    2 FORMAT (1H )
    3 FORMAT (1H+,3X,128A1)
      RETURN
      END
```

Exhibit G

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

MARCH 1980 VOLUME PAMI-2 NUMBER 2 (ISSN 0162-8828)

PUBLISHED BY THE IEEE COMPUTER SOCIETY

In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Group
Sonics and Ultrasonics Group
Systems, Man, and Cybernetics Society

PAPERS

CORRESPONDENCE

IEEE COMPUTER SOCIETY

The Computer Society is an association of people with professional interest in the field of computers. All members of the IEEE are eligible for membership in the Society upon payment of the annual Society membership fee of $8.00. Members of certain professional societies and other computer professionals are also eligible to be members of the Computer Society. For information on joining, write to IEEE Computer Society, P.O. Box 639, Silver Spring, MD 20901.

Officers

T. FENG, *President*
Dep. Comput. Sci.
Wright State Univ.
Dayton, OH 45435

O. N. GARCIA,
First Vice President
Systems Technology TIC
Dep. Elec. Eng. Syst.
Comput. Sci. Progr. (LIB 630)
University of South Florida
Tampa, FL 33620

M. E. SLOAN
Second Vice President
Chapters
Dep. Elec. Eng.
Michigan Technological University
Houghton, MI 49931

S. F. LUNDSTROM
Secretary
Burroughs Corp.
P.O. Box 517
Paoli, PA 19301

R. E. THEISEN
Treasurer
2667 Fitzhugh Rd.
Winter Park, FL 32792

Governing Board

Term Ending December 31, 1980

L. D. AMDAHL
K. R. ANDERSON
G. I. DAVIDA
R. HOELZEMAN
G. J. LIPOVSKI
S. E. MADNICK
R. L. RUSSO
J. SNYDER
H. S. STONE
R. E. THEISEN

Term Ending December 31, 1981

K. S. FU
P. L. HAZAN
P. ISAACSON
N. R. KORNFIELD
S. F. LUNDSTROM
E. A. PARRISH, JR.
A. POHM
R. RICE
R. G. STEWART
S. WINKLER

E. A. PARRISH, JR.
Vice President Software and Applications TIC

R. B. ARNDT
Vice President Conferences and Meetings

R. E. MERWIN
Vice President for Publications

Executive Secretary
H. HAYMAN
Box 639
Silver Spring, MD 20901

Junior Past President
M. G. SMITH
IBM T. J. Watson Res. Center
Yorktown Heights, NY 10598

Technical Committee Chairmen
Computer Architecture
T. AGERWALA
Computer Communications,
K. J. THURBER
Computer Elements,
J. ZASIO
Computer Packaging,
D. FRANK
Data Base Engineering, S. MADNICK
Design Automation,
W. M. VAN CLEEMPUT
Distributed Processing, C. R. VICK
Fault-Tolerant Computing,
A. L. HOPKINS
Machine Intelligence and Pattern Analysis,
Y. T. CHIEN
Mass Storage, I. TJOMSLAND
Mathematical Foundations of Computing, R. V. BOOK
Microprogramming, S. HUSSON
Mini/Microcomputers, P. I. HAZAN
Ocean Engineering, G. WILLIAMS
Operating Systems, H. L. APFELBAUM
Security and Privacy, G. I. DAVIDA
Simulation, N. SCHNEIDEWIND
Software Engineering,
B. BOEHM
Test Technology, K. R. ANDERSON
TIC—Software and Applications,
E. A. PARRISH, JR.
TIC—Systems Technology,
O. N. GARCIA

Standing Committee Chairmen
Admissions and Advancement,
E. A. PARRISH, JR.
Awards, R. RICE
Chapter Activities,
M. E. SLOAN
Area Committees
Mideastern, C. B. HENSLEY
Midwestern, W. H. HUEN
Northeastern, R. G. MATTESON
Ohio Valley, R. G. HOELZEMAN
Southeastern, W. D. CARROLL
Southwestern, W. K. KING
Western, R. NEILSON
Eastern Hemisphere and Latin America, R. C. BARQUIN
Chapter Development,
R. E. THEISEN
Distinguished Visitors Program,
E. M. REINGOLD
Newsletter, L. H. MAXSON
Student Activities Program,
C. SLIVINSKY
Tutorial Program, D. PESSEL
COMPCON
Spring, S. FERNBACH
Spring 1979, D. BROWN
Fall, D. HARTMANN
Fall 1980, M. D. ABRAMS
COMPSAC, S. S. YAU
COMPSAC 1980, R. E. MERWIN
Computer Standards,
R. G. STEWART
Conferences and Meetings,
R. B. ARNDT
Conference and Meetings Coordinator, G. J. LIPOVSKI
Constitution and Bylaws,
D. JACOBSOHN
Education
D. C. RINE
Fellows, S. LEVINE
Finance, M. E. SLOAN
Intersociety, M. G. SMITH
Membership and Transfers,
R. E. THEISEN
Nominations, M. G. SMITH
Publications, R. E. MERWIN
Student and Faculty Relations,
O. N. GARCIA

IEEE Division V Director
D. B. SIMMONS

Representatives to Other Organizations
AFIPS Directors, S. S. YAU,
R. B. ARNDT, T. FENG
AFIPS Executive Committee,
S. S. YAU
Annual Simulation Symposium,
O. N. GARCIA
IEEE Council on Oceanic Engineering,
G. N. WILLIAMS,
D. H. STOMBERG
IEEE Liaison with Information Theory Group, R. E. MILLER
IEEE Solid-State Circuits Council,
R. A. HENLE, F. K. BUELOW
Institute for Certification of Computer Professionals,
J. N. SNYDER, D. JACOBSOHN
Intersociety, M. G. SMITH
National Computer Conference Board, D. B. SIMMONS
Search Committee,
D. B. SIMMONS

THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

Officers

LEO YOUNG, *President*
C. LESTER HOGAN, *Executive Vice President*
BRUNO O. WEINSCHEL, *Secretary*
DONALD S. BRERETON, *Treasurer*
THEODORE H. BONN, *Vice President, Publication Activities*
RICHARD J. GOWEN, *Vice President, Professional Activities*
ROBERT E. LARSON, *Vice President, Technical Activities*
BENJAMIN J. LEON, *Vice President, Educational Activities*
LARRY K. WILSON, *Vice President, Regional Activities*
DICK B. SIMMONS, *Division V Director*

Headquarters Staff

ERIC HERZ, *Executive Director and General Manager*
RICHARD M. EMBERSON, *Executive Consultant*

THOMAS W. BARTLETT, *Controller*
DONALD CHRISTIANSEN, *Editor of Spectrum*
JOSEPH D. CODISPOTI, *Staff Director, Public Relations*
IVAN G. EASTON, *Staff Director, Standards*
LEO FANNING, *Staff Director, Professional Activities*
ELWOOD K. GANNETT, *Staff Director, Publishing Services*
PATRICIA LECH, *Staff Director, Field Services*
NEIL D. PUNDIT, *Staff Director, Technical Activities*
EMILY L. SIRJANE, *Staff Director, Corporate Services*
CHARLES F. STEWART, JR., *Staff Director, Administration Service*
JOHN F. WILHELM, *Staff Director, Educational Services*

Publications Department

H. JAMES CARTER, *Associate Staff Director*
Production Managers: ANN H. BURGMEYER, CAROLYNE ELENOWITZ*, GAIL S. FERENC, ISABEL NAREA
Supervisor, Special Publications: JOSEPH MORSICATO
Associate Editors: PRIJONO HARDJOWIROGO, PATRICIA H. NOLAN, KALLIE ZAPITI
**Responsible for this Transactions.*

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE is published bimonthly by The Institute of Electrical and Electronics Engineers, Inc. **Headquarters:** 345 East 47 Street, New York, NY 10017. Responsibility for the contents rests upon the authors and not upon the IEEE, the Society, or its members. **IEEE Service Center** (for orders, subscriptions, address changes): 445 Hoes Lane, Piscataway, NJ 00854. **Telephone:** Headquarters 212-644 + extension: Information-7900, General Manager-7910, Controller-7748, Education Services-7860, Publishing Services-7560, Regional/Section Services-7751, Standards-7960, Technical Services-7890, IEEE Service Center 201-981-0060, Washington Office/Professional Services 202-785-0017. **NY Telecopier:** 212-752-4929. **Telex:** 236-411 (International messages only). Individual copies: IEEE members $5.00 (first copy only), nonmembers $10.00 per copy. Annual subscription price: IEEE members, dues plus Society fee. Price for nonmembers on request. Available in microfiche and microfilm.

REFERENCES

[1] S. Watanabe, *Cognition and Pattern* (in Japanese). Tokyo, Japan: Iwanami, 1978.

[2] H. Bergson, *Évolution Créatrice*. Paris, France: Felix Alcan, 1907.

[3] S. Fujiwhara, *Jap. J. Astron. Geophys.*, vol. 5, p. 143, 1923; *J. Roy. Met. Soc.*, vol. 49, p. 89, 1923; see also S. Fujiwhara, *Kumo wo Tsukamu Hanashi* (in Japanese). Tokyo, Japan: Iwanami, 1940.

[4] N. Wiener, *Cybernetics*, 2nd ed. Cambridge, MA: MIT Press, 1962; see also *The Human Use of Human Being*, 2nd ed. Boston, MA: Houghton Mifflin, 1956.

[5] S. Watanabe, "The cybernetical view of time," in *Progress in Biocybernetics*, vol. 3, Wiener and Schade, Eds. Amsterdam, The Netherlands: Elsevier, 1966; see also "Nobert Wiener and cybernetical concept of time," *IEEE Trans. Syst., Man, Cybern.*, p. 372, May 1975.

[6] P. Glansdorff and I. Prigogine, *Thermodynamical Theory of Structure, Stability and Fluctuation*. New York: Wiley, 1971.

[7] S. Watanabe, "Information-theoretical aspects of inductive and deductive inference," *IBM J. Res. Develop.*, vol. 4, p. 208, 1960.

[8] —, "Learning process and inverse H-theories," *IRE Trans. Inform. Theory*, vol. IT-8, p. 248, 1962.

[9] —, *Knowing and Guessing*. New York: Wiley, 1969, p. 287.

[10] —, *Knowing and Guessing*. New York: Wiley, quoted in footnote 9, p. 395; "Karhunen–Loève expansion and factor analysis," in *Trans. 4th Prague Conf. on Inform. Theory, etc.*, 1965. Prague: Czechoslovak Acad. Sci., 1969.

[11] S. Watanabe and E. T. Harada, "A dynamical model of clustering," in *Proc. 2nd Int. Joint Conf. on Pattern Recognition*, Copenhagen, 1974, p. 413; S. Watanabe, "Further report on coalescence model of clustering," in *Proc. 3rd Int. Joint Conf. on Pattern Recognition*, San Diego, 1976, p. 176; "Application of dynamical coalescence model to pattern recognition," in *Proc. 4th Int. Joint Conf. on Pattern Recognition*, Kyoto, 1978.

[12] S. Watanabe, *Knowing and Guessing*. New York: Wiley, 1969, pp. 14, 17.

[13] A. Rosenfeld, R. A. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, no. 6, p. 420, 1976.

Digital Image Enhancement and Noise Filtering by Use of Local Statistics

JONG-SEN LEE

Abstract—Computational techniques involving contrast enhancement and noise filtering on two-dimensional image arrays are developed based on their local mean and variance. These algorithms are nonrecursive and do not require the use of any kind of transform. They share the same characteristics in that each pixel is processed independently. Consequently, this approach has an obvious advantage when used in real-time digital image processing applications and where a parallel processor can be used. For both the additive and multiplicative cases, the *a priori* mean and variance of each pixel is derived from its local mean and variance. Then, the minimum mean-square error estimator in its simplest form is applied to obtain the noise filtering algorithms. For multiplicative noise a statistical optimal linear approximation is made. Experimental results show that such an assumption yields a very effective filtering algorithm. Examples on images containing 256 × 256 pixels are given. Results show that in most cases the techniques developed in this paper are readily adaptable to real-time image processing.

Index Terms—Digital image enhancement, local statistics, noise filtering, real-time processing.

Manuscript received December 16, 1977; revised April 29, 1979.

The author is with the Naval Research Laboratory, Washington, DC 20375.

INTRODUCTION

Image processing on digital computers has been gaining in acceptance in recent years [1]–[3]. Early techniques in image processing concentrated mostly on procedures that were carried out computationally in the frequency domain (Fourier or Walsh), which was a natural extention of one-dimensional linear signal processing theory. In due course it became well-known that computing a two-dimensional transform for a large data array is a very time-consuming activity even with fast transform techniques on large computers. Hence, implementation of frequency domain procedures for real-time processing of images appears less than promising. More recent works based on an application of Kalman filtering algorithm [4] or Bayesian estimation extended to two-dimensional arrays led to the concept of a recursive filtering algorithm [5], [6]. The power of recursive algorithms for real-time one-dimensional signal processing are well established. However, when applied to a two-dimensional array, the algorithm operates in the spatial domain in which pixels have to be processed sequentially. As a consequence, the procedure is no longer computationally efficient and loses its attraction for real-time processing.

Algorithms developed in this paper share a particular characteristic in that each pixel can be processed separately without waiting for its neighboring pixels to be processed. This characteristic permits a direct implementation of these algorithms for real-time image processing. Applying local statistics to image processing is not a new idea. Ketcham *et al.* [7] used the entire local histogram for real-time image enhancement, and Wallis [8] applied local mean and variance to filter out scan line noise with striking results. This paper extends this family of algorithms to contrast enhancement and noise filtering. Both additive white noise and multiplicative white noise cases are considered. Most additive noise filtering approaches utilize the fast Fourier transform, convolution, or recursive algorithms. In the transform and convolution methods, the autocorrelation between pixels is either assumed or approximated, and in the recursive algorithm, a linear causal or semi-causal autoregressive image model is generally assumed. The techniques based on the use of local mean and variance described in this paper deviate from these approaches. The basic assumption is that the sample mean and variance of a pixel is equal to the local mean and variance of all pixels within a fixed range surrounding it. The validity of this assumption is debatable but so are most other statistical image representations encountered in the current practice. In the additive noise filtering case, the *a priori* mean (variance) of the estimated image is calculated as the difference between the mean (variance) of the noise corrupted image and the mean (variance) of the noise by itself. This technique is extended to include multiplicative noise filtering and also the case involving both multiplicative and additive noise. Although this simple approach may not have the mathematical elegance and sophistication of a few other techniques, our experimental results and those reported by Wallis [8] indicate that the local mean and variance technique is a very effective tool in both contrast stretching and noise filtering applications.

Let x_{ij} be the brightness of the pixel (i, j) in a two-dimensional $N \times N$ image. The local mean and variance are calculated over a $(2n + 1) \times (2m + 1)$ window. The local mean is defined as

$$m_{i,j} = \frac{1}{(2n+1)(2m+1)} \sum_{k=i-n}^{n+i} \sum_{l=j-m}^{m+j} x_{k,l}. \quad (1)$$

Similarly, the local variance is

$$v_{i,j} = \frac{1}{(2n+1)(2m+1)} \sum_{k=i-n}^{i+n} \sum_{l=j-m}^{j+m} (x_{k,l} - m_{i,j})^2. \quad (2)$$

In this paper, Section II is devoted to spatial contrast enhancement for which only the local mean is required. Section III deals with additive noise filtering. Section IV treats images corrupted by multiplicative noise, and Section V extends the results of Sections III and IV to images corrupted by both additive and multiplicative noise. In Section V, a simplified algorithm is discussed to approximate the local mean and variance and future research using local statistics is briefly mentioned.

II. Spatial Contrast Enhancement

Gray level rescaling, high-pass filtering, and histogram redistribution [9] are the most popular techniques in image contrast enhancement. Wallis [8] suggested an algorithm based on local mean and variance in which each pixel is required to have a "desirable" local mean m_d and a "desirable" local variance v_d such that

$$x'_{i,j} = m_d + \sqrt{\frac{v_d}{v_{i,j}}}\,(x_{i,j} - m_{i,j}) \quad (3)$$

where, in (3), $m_{i,j}$ and $v_{i,j}$ are local mean and variance as given by (1) and (2). It is easy to verify that the $x'_{i,j}$ has a mean m_d and variance v_d if we consider $m_{i,j}$ and $v_{i,j}$ as the true mean and variance of $x_{i,j}$. The main drawback of this technique is that it tends to enhance subtle details at the expense of the principal features which are lost in the process. Fig. 1 shows the original image and the image processed by (3). The river in the original image and other large objects are difficult to recognize in the processed image. To compensate for this, an algorithm is designed such that a pixel $x_{i,j}$ will maintain its local mean, and yet permits its variance to be modified by a factor of its original local variance. The modified algorithm is

$$x'_{i,j} = m_{i,j} + k(x_{i,j} - m_{i,j}) \quad (4)$$

where k, the gain, is the ratio of new local standard deviation to the original standard deviation. This scheme has a distinct computational advantage over the scheme rooted in (3). The computation of local variance $v_{i,j}$ is not required and only $m_{i,j}$ need be computed. If $k > 1$, the image will be sharpened as if acted upon by a high-pass filter. If $0 \leqslant k < 1$, the image will be smoothed as if passed through a low-pass filter. In the extreme case, $k = 0$ and $x'_{i,j}$ is equal to its local mean $m_{i,j}$.

This algorithm can be easily extended to enhance images with an ill-distributed histogram. Let $g(x)$ be the gray level rescaling function [9], then (4) is modified to

$$x'_{i,j} = g(m_{i,j}) + k(x_{i,j} - m_{i,j}). \quad (5)$$

In the case of a linear gray level stretch, $g(x)$ is a linear relation. Several images are processed and shown in Fig. 1 using $g(x) = ax + b$, where $a = 0.9$ and $b = 13$ to allow contrast enhancement at both ends of gray scale (between 0 and 255). The linear function $g(x)$ used in this picture yields an effective contrast stretch in both the highlights and the dark areas of the image. The window size of 3×3 or 5×5 is large enough to carry out contrast enhancement. For noise filtering (to be discussed in later sections), however, a 7×7 or higher dimensional window is more desirable but at the expense of image resolution. All images of Fig. 1 are processed by the use of a 5×5 window. Fig. 1(c) shows that for $k = 2$, both the gross features and subtle details are enhanced in the same proportion. The case $k = 1$ has no effect on the image. For $k = 0.5$, the image is blurred as if passed through a low-pass filter. For $k = 0$, the image is averaged over the 5×5 window.

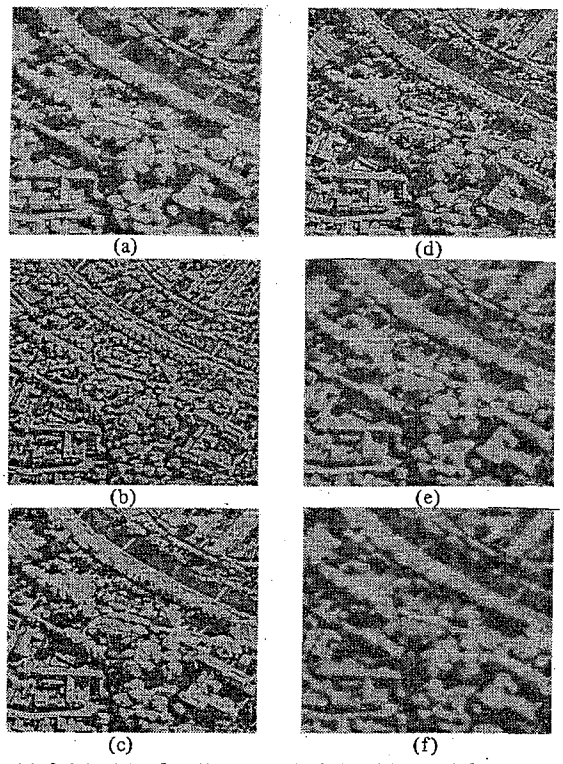


Fig. 1. (a) Original (or $k = 1$). (b) Wallis' algorithm. (c) $k = 2$. (d) $k = 3$. (e) $k = 0.5$. (f) $k = 0$.

III. Additive Noise Filtering

In this section, attention is focused on filtering of images degraded by white additive noise. Most current approaches to this problem employ frequency domain techniques, direct inversion, or recursive Kalman filtering. In the present paper, we derive a very simple algorithm based on the use of local mean and variance of an image. Let $z_{i,j}$ be the degraded pixel $x_{i,j}$; then

$$z_{i,j} = x_{i,j} + w_{i,j} \quad (6)$$

where $w_{i,j}$ is a white random sequence with

$$E[w_{i,j}] = 0$$

and $E[w_{i,j}w_{k,l}] = \sigma_1^2 \delta_{i,k}\delta_{j,l}$, where $\delta_{i,k}$ is the Kronecker delta function and E is the expectation operator. (The distribution of $w_{i,j}$ is irrelevant in this derivation.) The problem is to estimate $x_{i,j}$, given $z_{i,j}$ and the noise statistics.

From (6), we have

$$\bar{x}_{i,j} \triangleq E[x_{i,j}] = E[z_{i,j}] = \bar{z}_{i,j} \quad (7)$$

and

$$Q_{i,j} \triangleq E[(x_{i,j} - \bar{x}_{i,j})^2] = E[(z_{i,j} - \bar{z}_{i,j})^2] - \sigma_1^2. \quad (8)$$

In most filtering techniques, the *a priori* mean and variance of $x_{i,j}$ are derived from an assumed autocorrelation model or, recursively, from an autoregressive Markov sequence. In either method, it is an approximation. Assume that $\bar{x}_{i,j}$ and $Q_{i,j}$ are the *a priori* mean and variance of $x_{i,j}$, which in turn are approximated by the local mean and variance of (7) and (8). Under this assumption, it is very easy to obtain a filtering algorithm either by minimizing the mean-square error or

by weighted least-square estimation [10]. Both methods will produce the same algorithm. The estimated $x_{i,j}$, $\hat{x}_{i,j}$ is computed by

$$\hat{x}_{i,j} = \bar{x}_{i,j} + k_{i,j}(z_{i,j} - \bar{x}_{i,j}) \tag{9}$$

where the gain is given by

$$k_{i,j} = \frac{Q_{i,j}}{Q_{i,j} + \sigma_1^2}. \tag{10}$$

Equation (9) can be written as

$$\hat{x}_{i,j} = (1 - k_{i,j})\bar{x}_{i,j} + k_{i,j}(z_{i,j}). \tag{11}$$

Since $Q_{i,j}$ and σ_1^2 are both positive, $k_{i,j}$ will lie between 0 and 1. A simple intuitive interpretation is that for a low signal-to-noise ratio region $Q_{i,j}$ is small compared with σ_1^2, $k_{i,j} \simeq 0$, and the estimated $\hat{x}_{i,j}$ is $\bar{x}_{i,j}$. Conversely, for a high signal-to-noise ratio region, $Q_{i,j}$ is much larger than σ_1^2, $k_{i,j} \simeq 1$, and $\hat{x}_{i,j} \simeq z_{i,j}$, the corrupted pixel. The use of different window sizes will greatly affect the quality of processed images. If the window is too small, the noise filtering algorithm is not effective. If the window is too large, subtle details of the image will be lost in the filtering process. Our experiments indicate that a 7×7 window is a fairly good choice. All images presented in this and later sections are processed by the 7×7 window.

Fig. 2 shows the original image, the image contaminated with additive noise and the estimated image produced by the local mean and variance algorithm. Clearly, in a smooth area, the pixel is averaged over the window and in a high contrast or edge area, the noise corrupted pixels are weighted higher than their local mean value. Fig. 2(d), (e), and (f) are the plots of intensity along a specific scan line for the original, the noise corrupted, and the processed images, respectively. This algorithm works equally well for an image corrupted by a Gaussian noise. Results for the latter case are shown in Fig. 3(a) and (b).

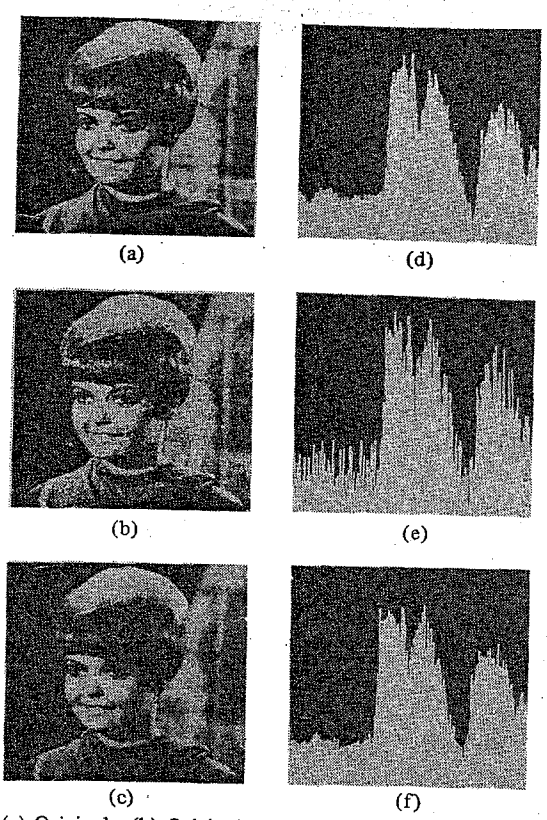


Fig. 2. (a) Original. (b) Original plus additive uniform noise (−30, 30). (c) Additive noise removed with (7×7) mesh, $\sigma_1^2 = 300$. (d) Original intensity profile along a scan line. (e) Profile at (d) contaminated by additive noise. (f) Profile at (e) filtered for noise.

Fig. 3. (a) Original plus additive Gaussian noise, $\sigma_1^2 = 300$. (b) Noise removed.

IV. Multiplicative Noise Filtering

Images containing multiplicative noise have the characteristic that the brighter the area the noisier it is. Mathematically, the degraded pixel can be represented by

$$z_{i,j} = x_{i,j}u_{i,j} \tag{12}$$

where $E[u_{i,j}] = \bar{u}_{i,j}$ and

$$E[(u_{i,j} - \bar{u}_{i,j})(u_{k,l} - \bar{u}_{k,l})] = \sigma_2^2 \delta_{i,k}\delta_{j,l}.$$

Nahi and Naraghi [11] treat this problem via the Kalman–Bucy approach which necessitates solving nonlinear estimation problem by numerical integration. In this paper, the nonlinearity in (12) is treated differently. An optimal linear approximation of (12) is used to produce a filtering algorithm similar to that for the additive noise case. Experimental results show that the derived algorithm is a very promising one. Let

$$z'_{i,j} = A\, x_{i,j} + B\, u_{i,j} + C \tag{13}$$

where A, B, and C are nonrandom variables. They are to be chosen to minimize the mean-square error between $z_{i,j}$ and $z'_{i,j}$ and also to make $z'_{i,j}$ an unbiased estimate of $z_{i,j}$. For $z'_{i,j}$ to be unbiased estimate of $z_{i,j}$, we must have

$$A\,\bar{x}_{i,j} + B\,\bar{u}_{i,j} + C = \bar{x}_{i,j}\,\bar{u}_{i,j}$$

or

$$C = \bar{x}_{i,j}\,\bar{u}_{i,j} - A\,\bar{x}_{i,j} - B\,\bar{u}_{i,j}. \tag{14}$$

Substituting (14) into (13) and forming the mean-square error, we arrive at the performance index to be minimized,

$$J = E[A(x_{i,j} - \bar{x}_{i,j}) + B(u_{i,j} - \bar{u}_{i,j}) - (x_{i,j}u_{i,j} - \bar{x}_{i,j}\,\bar{u}_{i,j})]^2.$$

Upon carrying out the necessary mathematical procedures, we obtain the following relation:

$$z_{i,j} = \bar{u}_{i,j}\,x_{i,j} + \bar{x}_{i,j}(u_{i,j} - \bar{u}_{i,j}). \tag{15}$$

It is not surprising to find that (15) actually is the first-order Taylor series expansion of $z_{i,j}$ about $(\bar{x}_{i,j}, \bar{u}_{i,j})$.

The *a priori* mean and variance of $x_{i,j}$ are computed from (12) and are given by

$$\bar{x}_{i,j} = \bar{z}_{i,j}/\bar{u}_{i,j} \tag{16}$$

and

$$Q_{i,j} = \frac{\mathrm{var}(z_{i,j}) + \bar{z}_{i,j}^2}{\sigma_2^2 + \bar{u}_{i,j}^2} - \bar{x}_{i,j}^2 \tag{17}$$

in which $\mathrm{var}(z_{i,j})$ is the variance of $z_{i,j}$. The quantities $\bar{z}_{i,j}$ and $\mathrm{var}(z_{i,j})$ are approximated by the local mean and local variance of the corrupted image. Using (16) and (17), and

applying the Kalman filtering algorithm to (15), we have

$$\hat{x}_{i,j} = \bar{x}_{i,j} + k_{i,j}(z_{i,j} - \bar{u}_{i,j}\,\bar{x}_{i,j}), \tag{18}$$

in which

$$k_{i,j} = \frac{\bar{u}_{i,j}\,Q_{i,j}}{\bar{x}^2_{i,j}\,\sigma_2^2 + \bar{u}^2_{i,j}\,Q_{i,j}}. \tag{19}$$

An experimental example is shown in Fig. 4, in which the original image is corrupted by multiplicative noise uniformly distributed between 0.7 and 1.0, and the estimated image has been processed by the algorithm developed in this section. Considerable improvement is shown in the processed image, thus substantiating the effectiveness of the local mean and variance technique.

V. Filtering of Combined Additive and Multiplicative Noise

It is very easy to extend the algorithms of previous sections to deal with images corrupted by both additive and multiplicative noise. A noise-corrupted image is described by

$$z_{i,j} = x_{i,j}\,u_{i,j} + w_{i,j} \tag{20}$$

in which the statistical characteristics are the same as given in Sections III and IV. Assume that $u_{i,j}$ and $w_{i,j}$ are independent white noises. This independence assumption can be removed, but the result is a more complicated formulation. Following the idea of an optimal linear approximation of Section IV, we have

$$z'_{i,j} = \bar{u}_{i,j}\,x_{i,j} + \bar{x}_{i,j}(u_{i,j} - \bar{u}_{i,j}) + w_{i,j}.$$

The formulas for the *a priori* mean and variance of $x_{i,j}$ of Section IV are modified to read

$$\bar{x}_{i,j} = (\bar{z}_{i,j} - \bar{w}_{i,j})/\bar{u}_{i,j} \tag{21}$$

and

$$Q_{i,j} = \frac{\operatorname{var}(z_{i,j}) + \bar{z}^2_{i,j}}{\sigma_2^2 + \bar{u}^2_{i,j}} - \bar{x}^2_{i,j} - \sigma_1^2.$$

The filtering algorithm is

$$\hat{x}_{i,j} = \bar{x}_{i,j} + k_{i,j}(z_{i,j} - \bar{u}_{i,j}\,\bar{x}_{i,j} - \bar{w}_{i,j}) \tag{22}$$

in which

$$k_{i,j} = \frac{\bar{u}_{i,j}\,Q_{i,j}}{\bar{x}^2_{i,j}\,\sigma_2^2 + \bar{u}^2_{i,j}\,Q_{i,j} + \sigma_1^2}. \tag{23}$$

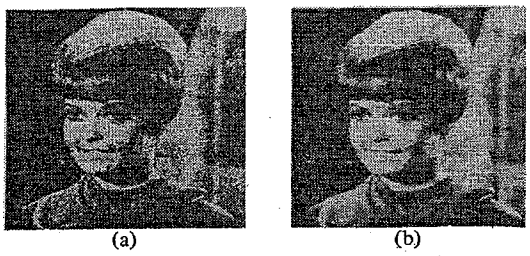
Fig. 5(a) shows the image corrupted by an additive noise uniformly distributed between gray levels −20 and +20 and also a multiplicative noise uniformly distributed between multiplicative factors 0.7 and 1.0. The processed image, Fig. 5(b) shows a very significant improvement over the original image.

VII. Remarks and Conclusions

The principal computational load of the developed algorithms is in the calculation of the local means and variances, especially the latter. To lighten this burden Wallis [8] proposed a fast algorithm in which the image is partitioned into square subregions over which the local mean and variance are computed. Then the local mean and variance of a particular pixel are approximated by the use of two-dimensional interpolation formulas. Results, as reported by Wallis [8], are quite favorable. It is believed that Wallis' approach will yield an equally impressive improvement when applied to our contrast enhancement and noise filtering algorithms.

(a) (b)

Fig. 4. (a) Multiplicative noise, $U(0.7, 1.0)$. (b) Multiplicative noise removed.



(a) (b)

Fig. 5. (a) Image corrupted by additive and multiplicative noise. (b) Restored.

In conclusion, image processing algorithms presented in this paper, based on considerations of the local image statistics, have a structure which makes them naturally suitable for parallel processing. Since the latter approach offers great computational economy, real or near real-time processing can be achieved. Future research in this area is to extend the method to image restoration of motion blur and other degradations characterized by local correlations around pixels.

Acknowledgment

The author wishes to thank Drs. I. Jurkevich and A. F. Petty for many helpful discussions and encouragements.

References

[1] T. S. Huang, Ed., *Picture Processing and Digital Filtering (Topics in Applied Physics)*, vol. 6. New York: Springer-Verlag, 1975.
[2] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
[3] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. New York: Academic, Aug. 1976.
[4] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Eng.*, ser. D, vol. 82, pp. 35-45, 1960.
[5] A. K. Jain, "A semicausal model for recursive filtering of two-dimensional images," *IEEE Trans. Comput.*, vol. C-26, Apr. 1977.
[6] N. E. Nahi and T. Assefi, "Baysian recursive image estimation," *IEEE Trans. Comput.*, vol. C-12, pp. 734–738, July 1972.
[7] P. J. Ketcham, R. W. Lowe, and J. W. Weber, "Real time image enhancement techniques," in *Proc. Seminar in Image Processing*, Pacific Grove, CA, Feb. 1976.
[8] R. Wallis, "An approach to the space variant restoration and enhancement of images," in *Proc. Symp. on Current Mathematical Problems in Image Science*, Naval Postgraduate School, Monterey, CA, Nov. 1976.
[9] W. Frei, "Image enhancement by histogram hyperbolization," *Comput. Graphics Image Processing*, vol. 6, pp. 286–294, June 1977.
[10] A. E. Bryson and Y. C. Ho, *Applied Optimal Control*. Waltham, MA: Blaisdell, 1969.
[11] N. E. Nahi and M. Naraghi, "A general image estimation algorighm applicable to multiplicative and non-Gaussian noise," Univ. of Southern California, Image Processing Inst. Semi-Annual Tech. Rep. 620, Sept. 30, 1975.

Exhibit H

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

NOVEMBER 1981 VOLUME PAMI-3 NUMBER 6 (ISSN 0162-8828)

PUBLISHED BY THE IEEE COMPUTER SOCIETY

In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Group
Sonics and Ultrasonics Group
Systems, Man, and Cybernetics Society

PAPERS

CORRESPONDENCE

BOOK REVIEWS

[5] R. E. Warmack and R. C. Gonzalez, "An algorithm for the optimal solution of linear inequalities and its application to pattern recognition," *IEEE Trans. Comput.*, vol. C-22, no. 12, pp. 1065–1075, 1973.
[6] A. Miyake and S. Shinmura, "An algorithm for the optimal linear discriminant functions," in *Proc. Int. Conf. Cybern. Soc.*, vol. 2, 1978, pp. 1447–1450.
[7] S. Shinmura and A. Miyake, "Optimal linear discriminant functions and their application," in *Proc. Comput. Software Appl. Conf.*, Chicago, IL, Nov. 6–8, 1979, pp. 167–172.
[8] A. Miyake, "Mathematical aspects of optimal linear discriminant functions," in *Proc. Comput. Software Appl. Conf.*, Chicago, IL, Nov. 6–8, 1979, pp. 161–166.
[9] H. Mengert, "Solution of linear inequalities," *IEEE Trans. Comput.*, vol. C-19, no. 2, pp. 124–131, 1970.
[10] F. W. Smith, "Pattern classifier design by linear programming," *IEEE Trans. Comput.*, vol. C-17, no. 4, pp. 367–372, 1968.
[11] R. G. Grinold, "Comment on 'Pattern classifier design by linear programming,'" *IEEE Trans. Comput.*, vol. C-18, no. 4, pp. 378–379, 1969.
[12] E. W. Cheney, *Introduction to Approximation Theory*. New York: McGraw-Hill, 1966.
[13] V. Klee, "What is a convex set," *Amer. Math. Mon.*, vol. 78, pp. 617–631, June–July, 1971.
[14] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1970.
[15] T. M. Cover, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," *IEEE Trans. Electron. Comput.*, vol. EC-14, no. 6, pp. 326–334, 1965.
[16] J. K. Bryan and D. L. Tebbe, "Generation of multivariate Gaussian data," Dep. Elec. Eng., Univ. Missouri, Colombia, Tech. Rep., Apr. 1970.
[17] G. N. Wassel, "Training a linear classifier to optimize the error probability," Ph.D. dissertation, Univ. California, Irvine, Dec. 1972.
[18] H. Do-Tu and M. Installe, "Learning algorithms for nonparametric solution to the minimum-error classification problem," *IEEE Trans. Comput.*, vol. C-27, pp. 648–659, 1978.

D. C. Clark received the B.S. degree in mathematics from the California Institute of Technology, Pasadena, in 1963, the Ph.D. degree in mathematics from Stanford University, Stanford, CA, in 1968, and the M.S. degree in computer science from the University of Tennessee, Knoxville, in 1980.

He has taught in the Departments of Mathematics at Rutgers University, Newark, NJ, and the University of Puerto Rico, Mayaguez. Currently, he is with the Los Angeles office of Pattern Analysis and Recognition Corporation.

R. C. Gonzalez (S'65–M'70) was born in Havana, Cuba, on August 26, 1942. He received the B.S.E.E. degree from the University of Miami, Coral Gables, FL, in 1965 and the M.E. and Ph.D. degrees in electrical engineering from the University of Florida, Gainesville, in 1967 and 1970, respectively.

He has been affiliated with the GT&E Corporation, the Center for Information Research at the University of Florida, NASA, and is presently IBM Professor of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville. He is a frequent consultant to industry and government in the areas of pattern recognition, image processing, and machine learning. He received the 1978 UTK Chancellor's Research Scholar Award, the 1980 Magnavox Engineering Professor Award, and the 1980 M. E. Brooks Distinguished Professor Award for his work in these fields. He is coauthor of the books *Pattern Recognition Principles*, *Digital Image Processing*, and *Syntactic Pattern Recognition: An Introduction*, all published by Addison-Wesley. He is an Associate Editor for the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS and the *International Journal of Computer and Information Sciences*.

Dr. Gonzalez is a member of several professional and honorary societies, including Tau Beta Pi, Phi Kappa Phi, Eta Kappa Nu, and Sigma Xi.

Real-Time Adaptive Contrast Enhancement

PATRENAHALLI M. NARENDRA, MEMBER, IEEE, AND ROBERT C. FITCH, MEMBER, IEEE

Abstract—A recursive filter approach is introduced to simplify real-time implementation of an adaptive contrast enhancement scheme for imaging sensors. With this scheme, even scenes possessing large global dynamic ranges (>40 dB) can be accommodated by the limited dynamic range (20 dB) of a display without losing the local contrast essential for image interpretation. This paper describes the recursive filter implementation of the local area contrast enhancement scheme using charge-coupled devices and the resultant real-time hardware capable of processing standard 525 and 875 line TV compatible video (from vidicons, videotape recorders, etc). Several examples from video imagery are included to demonstrate its effectiveness.

Index Terms—AGC, contrast enhancement, digital filter, dynamic range compression, FLIR, image enhancement, real-time processing, signal processing architectures, TV, video.

Manuscript received March 10, 1980; revised February 6, 1981. This work was supported in part by the U.S. Army Night Vision and Electro Optics Laboratory under Contract DAAG53-76-C-0195.

The authors are with the Systems and Research Center, Honeywell, Inc., Minneapolis, MN 55413.

I. INTRODUCTION

THE SCENE dynamic ranges encountered by imaging sensors can be much higher than the CRT display luminance range (~20 dB). Moreover, the scene intensity extremes are changing in time. It is obvious that some form of global automatic gain/bias control (AGC) system can be done by a linear scaling of the intensities governed by the scene extrema or the global scene variance.

But mere global gain and bias control does not suffice when the scene has a large dynamic range. When the high dynamic range scene information is squeezed into the limited range of a typical display, low contrast local details are not perceived. Therefore, in addition to the global gain and bias control we need local area contrast enhancement—to enhance the local contrast while compressing the global scene dynamic range presented to the display. Previous work in this direction included Ketcham [1] and Harris [2]. Both are essentially high

0162-8828/81/1100-0655$00.75

frequency emphasis schemes, and use local area statistics (mean and variance) computed on a sliding window to control the contrast and bias level at each image point. Note that these are local transformation schemes as distinct from global gray scale transformations also suggested for enhancing display contrast [3], [4].

Our present approach replaces the complex computations of the local area statistics with simple two-dimensional recursive low-pass filters. This enables a very simple real-time implementation at video rates. CCD line delays are used to provide the time delays needed in the recursive filter. This in turn results in a sampled analog structure that uses CCD line delays, analog multipliers, and adders in the 10–30 MHz rates encountered in real-time video applications.

In Sections II and III we describe the basic local area contrast enhancement (LACE) scheme, discuss various alternative implementations of the local area statistics computation, and discuss the application of the recursive filter to greatly simplify these operations. The description of the simple hardware implementing this approach follows in Section IV. We include several examples of videotaped imagery enhanced by this real-time hardware to demonstrate its effectiveness.

II. The Local Area Contrast Enhancement (LACE) Scheme

Fig. 1 is a one-dimensional illustration of a scene which can be seen as a slowly varying background envelope. Superimposed upon it is higher frequency local variation representing scene detail and other information cues. The problem with real-life imagery is that the gain has to be made small enough to squeeze the wide dynamic range of the scene into the small (20:1) luminance range of the display. Then the *local* contrast (i.e., difference in intensities) falls below the contrast sensitivity threshold ΔI of the human eye and will not be perceived. The contrast sensitivity threshold ΔI is a function of both the local-area size and the average local intensity. For a given local-area size, the contrast sensitivity obeys Weber's law (i.e., is proportional to the intensity, $\Delta I/I = C_s$, a constant). This contrast C_s can vary from less than 1 percent for large areas to infinity for very small local areas.

Local area contrast enhancement (LACE) schemes operate under the premise that the large slowly varying intensity excursions can be reduced without degrading the information in the scene. The local contrast can then be enhanced (by increasing the local gain) without exceeding the dynamic range of the display. Fig. 1(b) illustrates this; the low frequency envelope is brought closer to the global mean while the higher frequency local variations are now amplified above the contrast sensitivity threshold ΔI. This does not significantly degrade the information content of the image since the relative brightness of distant areas within the scene contributes little to discrimination.

A LACE formulation that addresses the above visual psychophysical considerations is derived here. It performs the following functions:

- vary the local average brightness (bias) so that overall dynamic range of scene is compressed;
- enhance local variations above the contrast sensitivity threshold of the human eye; and

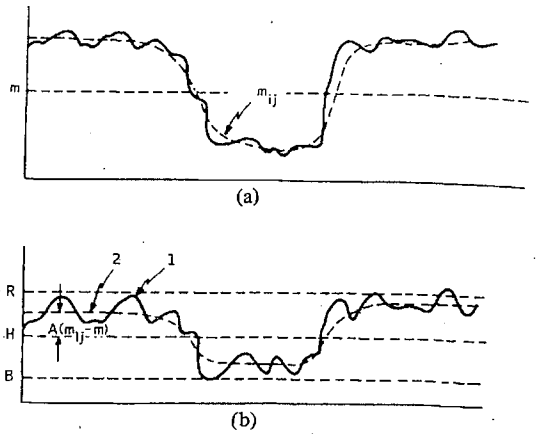


Fig. 1. (a) Original one-dimensional scene example. (b) LACE applied: 1) high frequency component and 2) low frequency component.

- automatically fit the intensity extremes in the enhanced scene to the display limits.

A functional description of the algorithm is shown in Fig. 2. The image intensity at each point is transformed based on local area statistics—the local mean M_{ij} and the local standard deviation σ_{ij} computed on a local area surrounding the point. The transformed intensity is then

$$\hat{I}_{ij} = G_{ij}(I_{ij} - M_{ij}) + M_{ij} \tag{1}$$

where the local gain is

$$G_{ij} = \alpha \frac{M}{\sigma_{ij}}, \qquad 0 < \alpha < 1$$

where M is the global mean.

The local area mean is first subtracted from the image at every point. A variable gain is applied to the difference to amplify the local variations. A portion of the local mean M_{ij} is then added back to restore the subjective quality of image. The local gain G_{ij} is itself locally adaptive, being proportional to M_{ij}, to satisfy psychovisual considerations (Weber's law); and inversely proportional to σ_{ij}, so that areas with small local variance receive larger gain. To prevent the gain from being inordinately large in areas with large mean and small standard deviation, the local gain is actually controlled as in Fig. 3.

Note that this formulation is virtually identical to those of [1] and [2]. The main contribution of this paper is in an efficient recursive architecture for the implementation of this scheme, discussed in the next section.

III. Real-Time Implementation

Three alternate implementations have been considered to realize the basic scheme in Fig. 2. They differ mainly in the way the local area mean and standard deviation are computed in real time.

Nonrecursive Implementation

The most obvious implementation is to have an $M \times M$ sliding window using shift registers (line delays) to perform nonrecursive computation of the mean and standard deviation [1]. This requires simultaneous access to M lines—which implies an inordinate amount of line delays and high-speed addition (over $2M$ adds/pixel at 8–30M samples/s) at the video rates en-

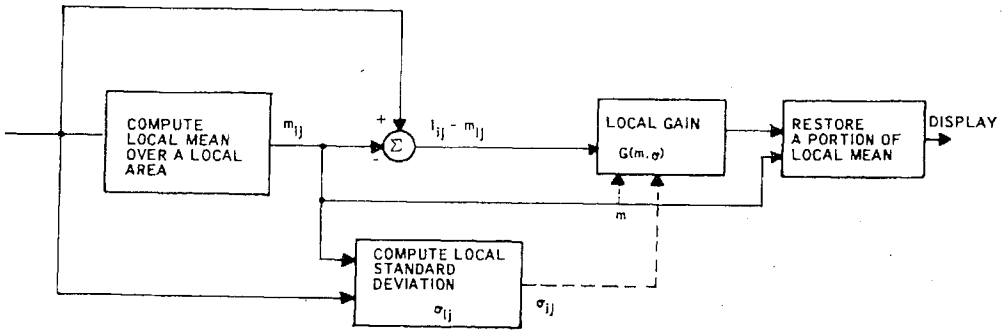


Fig. 2. Functional flow description of the local area contrast enhancement algorithm.

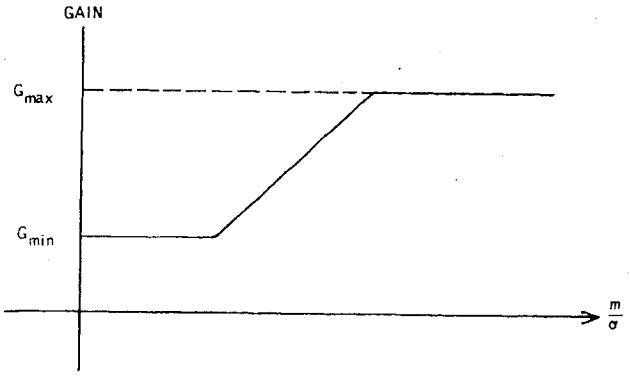


Fig. 3. Local area gain curve to prevent excessive gain variations.

countered in imaging sensors. A 15 × 15 window requires real-time buffering (in the shift registers) of 15 lines of video, and 225 adds/sample period.

Bilinear Interpolation

To reduce the hardware complexity of nonrecursive computation on a sliding window, an alternate approach was computer simulated. Here, the local mean and standard deviation are computed on *nonoverlapping* local areas of size M and stored for each frame. The local mean and standard deviation at any point in the next frame are then estimated by a simpler bilinear interpolation. This approach is effective because it cuts down the mean and standard deviation computation to N^2 MAD's versus M^2N^2 (a saving by a factor of over 100) without any of the grid-like artifacts that would be inherent in other nonoverlapping window approaches. This is because the interpolated values of the mean standard deviation vary smoothly across the region boundaries.

This approach was found to be very efficient and effective in computer simulations. But its real-time implementation is still not as simple as we would like it to be. Although the number of multiplies and adds per pixel is greatly reduced, the storing, addressing, and updating of the local area statistics and the interpolation require a complex architecture.

Recursive Implementation

The recursive approach achieves a very simple structure suitable for real time implementation. It is obvious that the local mean and standard deviation computation is in fact nonrecursive low-pass filtering of the input and (input - local mean)^2 with a rect function. If we replace these nonrecursive low-pass filters by two-dimensional recursive low-pass filters, we stand to gain considerable simplicity in the resultant implementation.

Fig. 4 is a realization of the LACE scheme using linear recursive low-pass filters to estimate the local mean and standard deviation. Here we replace the local average function with an equivalent two-dimensional recursive low-pass filter. The local standard deviation σ_{ij} is approximated by a similarly low-passed version of the absolute difference between the image intensity I_{ij} and the local mean estimate M_{ij}.[1]

A two-dimensional separable first-order recursive low-pass filter is the basic building block in Fig. 4. This filter has a frequency response given by the product of the two one-dimensional filter responses:

$$|H(f_x, f_y)| = [1 + (f_x/f_c)^2]^{-1/2} [1 + (f_y/f_c)^2]^{-1/2} \quad (3)$$

where f_c is the 3 dB cutoff frequency of the low-pass filter. The equivalent sampled data filter has a two-dimensional Z transform,

$$H(z_1, z_2) = \frac{\gamma^2}{(1 - e^{-\gamma} z_1^{-1})(1 - e^{-\gamma} z_2^{-1})} \quad (4)$$

where

$$\gamma = 2\pi f_c/f_s.$$

Changing γ changes the effective size of the local area, i.e., the area over which the local mean averaging is done. The above separable filter can be realized using two distinct (but functionally equivalent) structures.

Nonseparable Implementation: In the Z domain, let the output be $Y(z_1, z_2)$ and the input be $X(z_1, z_2)$. Then,

$$\frac{Y(z_1, z_2)}{X(z_1, z_2)} = \frac{\gamma^2}{(1 - e^{-\gamma} z_1^{-1})(1 - e^{-\gamma} z_2^{-1})}. \quad (5)$$

This form, implemented directly, gives the recursive relation

$$y(m, n) = \gamma^2 x(m, n) + e^{-\gamma} y(m - 1, n) + e^{-\gamma} y(m, n - 1) - e^{-2\gamma} y(m - 1, n - 1) \quad (6)$$

where m is the row number and n is the column number in the image.

[1] Note that for a Gaussian distribution, $E|x - \mu| = \sqrt{2/\pi}\ \sigma$, where μ is the mean and σ is the standard deviation.

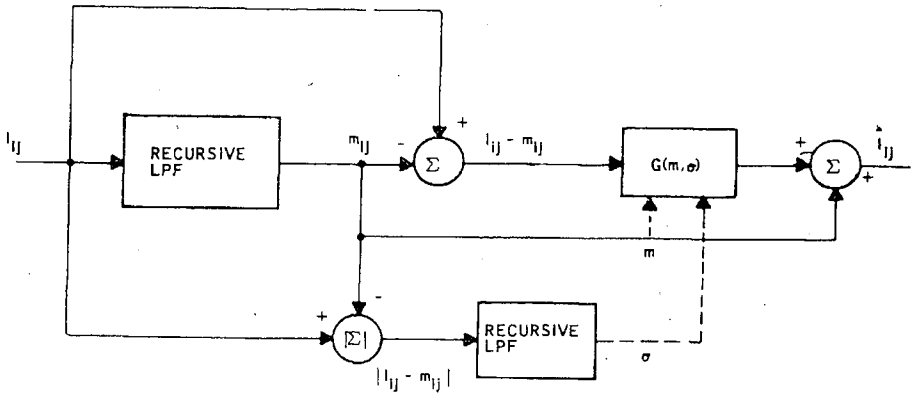


Fig. 4. LACE with linear recursive low-pass filters.

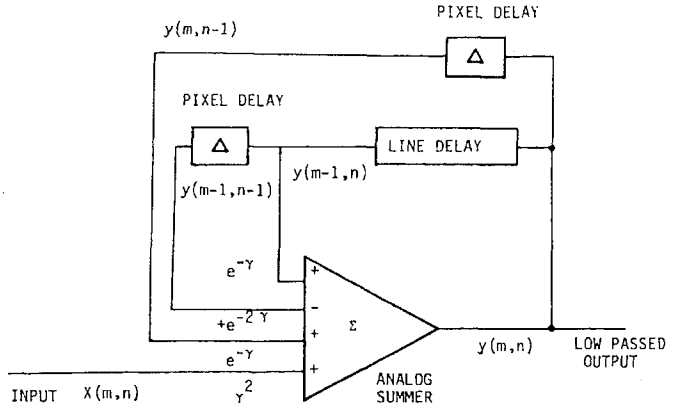


Fig. 5. Implementation of two-dimensional recursive low-pass filter (Approach 1).

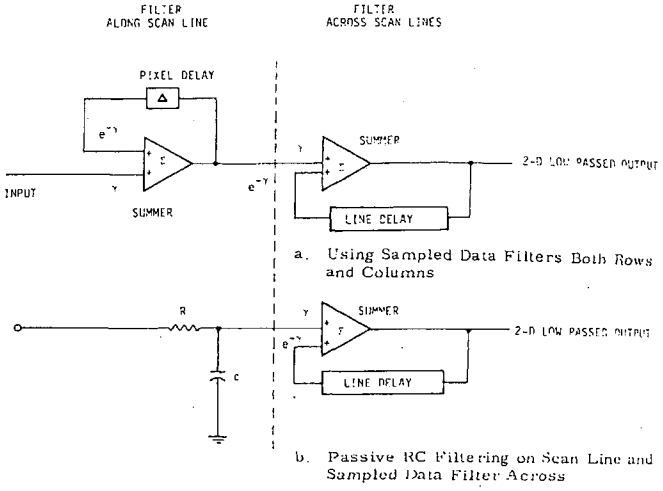


Fig. 6. Separable implementation of two-dimensional low-pass filter (Approach 2).

The schematic for implementing this realization on a real-time video stream is shown in Fig. 5. We see that just one line delay and the two pixel delays are necessary to perform the filtering.

Separable Implementation: Define a new intermediate variable

$$W(z_1, z_2) = \frac{\gamma}{(1 - e^{-\gamma} z_1^{-1})} \cdot X(z_1, z_2). \tag{7}$$

Then,

$$\frac{Y(z_1, z_2)}{X(z_1, z_2)} = W(z_1, z_2) \cdot \frac{\gamma}{(1 - e^{-\gamma} z_2^{-1})}. \tag{8}$$

Therefore,

$$y(m, n) = \gamma w(m, n) + e^{-\gamma} y(m - 1, n) \tag{9}$$

and

$$w(m, n) = \gamma x(m, n) + e^{-\gamma} w(m, n - 1). \tag{10}$$

Thus we break up the two-dimensional filter into two one-dimensional filters in cascade. Fig. 6 shows this filter realized in the above manner. The output of this filter is exactly equivalent to that of Fig. 5 because the transfer function is separable. This realization has certain advantages over the nonseparable implementation as described below.

The parameter γ controls the effective size of the local area. For local area (impulse response area) sizes desired in LACE, γ is in the range 0.1-0.3, which makes $e^{-\gamma} \cong 1 - \gamma$, in the range 0.9-0.7. In the nonseparable realization, the smallest weight of the summing amplifier, $\gamma^2 = 0.1$-0.01, is much smaller than the largest weight $1 - \gamma = 0.7$-0.9. The coefficient precision needed in the summer in Fig. 5 is therefore very stringent. On the other hand, the coefficient ranges in the separable formulation is much smaller: $\gamma = 0.1$-0.3 for the smallest weight and $1 - \gamma = 0.9$-0.7 for the largest. Therefore, the separable structure is less susceptible to charge-coupled device (CCD) noise and amplifier gain variation than is the corresponding nonseparable structure. In a corresponding digital implementation as well, the separable realization would require a smaller precision adder.

The separable structure has an additional advantage in an analog implementation. The first low-pass filter (along the scan direction) can be easily implemented in a passive RC first-order circuit, as shown in Fig. 6(b). At video frequencies, the required resistor (R) and capacitor (C) values are very reasonable. This eliminates the need for the single pixel delay for the

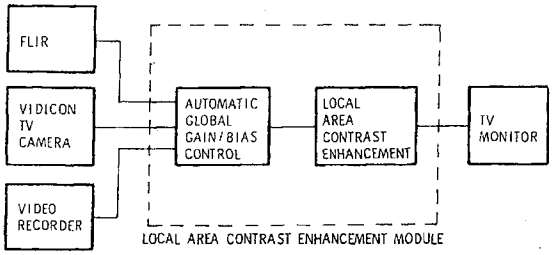


Fig. 7. Local area contrast enhancement module configuration.

first filter. However, the second low-pass filter in the vertical direction is still a sampled data filter as before, requiring one line delay.

We see that the first-order recursive filter above can achieve equivalent size local averaging as a nonrecursive window with just one line delay and two weighting coefficients. The impulse response is an exponential, $h(m, n) = e^{-(m+n)\gamma}, m, n \geq 0$ controls the size of the impulse response. Making γ smaller makes the impulse response wider.

In the hardware description that follows, we used $\gamma = 0.1$. Note also that since the impulse response of the recursive filter is causal, the local area mean at a point evaluated by the filter is a function only of the upper left quadrant centered at that point. Therefore, the resultant filter is not zero-phase as a nonrecursive filter with a sliding window would be. Despite this, in the simulations we have performed and in the results of the real-time hardware, we observed no undue distortion caused by the nonzero phase of the recursive filter. Also since the filter is causal, we do not have to delay the input video to be in phase with the local mean before subtracting the local mean to Fig. 4. These advantages outweigh the nonzero phase property of the filter.

IV. Local Area Contrast Enhancement Hardware

The local area contrast enhancement uses two identical two-dimensional recursive filters—one to compute the local mean and the other to compute the local standard deviation as the schematic in Fig. 4 showed. The separable realization of the filter in Fig. 6, being the simplest, was chosen for hardware implementation. Commercially available CCD video line delays were used in the two recursive filters. All the remaining circuitry to realize the various summing and variable gain functions in Fig. 4, employ conventional high bandwidth analog op amps and multipliers. This breadboard was designed and built entirely with off-the-shelf components. Inspite of this constraint, the simplicity achieved by the recursive sampled analog filter approach resulted in a compact (6 X 6 in card) implementation of the entire LACE scheme.

Fig. 7 shows the general configurations in which this hardware can be used. It can receive 525 line or 875 line TV compatible video from any source—an FLIR, videotape recorder, video disk recorder, vidicon, etc. The output is a composite video signal capable of driving a 75 Ω load.

The input to the LACE unit is first automatically scaled between 0 and 1 V by the global gain/bias control unit using the frame video mean and standard deviation. This adjusts the contrast and brightness on a sync separated video input from a video disk, tape, camera, etc., so that effective use of the CCD dynamic ranges is made in the subsequent stages.

This module has been extensively tested with real-time videotaped 525 line imagery (at 30 frames/s). A few examples from this evaluation are reproduced in Fig. 8. These are photographs taken off the display before and after enhancement.

In the interests of a fair trial, the contrast and brightness levels on the monitors were not changed for the two conditions. In both cases, the video into the display received global scaling (i.e., had the same global extrema). This is in fact evident in the photographs themselves—the *global* contrast and average brightness remain the same before and after enhancement.

The most dramatic improvement in local area contrast can be seen in the FLIR images in Fig. 8. The large dark areas consist of smoke which tends to drive the video blacker than black on a conventional display, obscuring surrounding details. Local area contrast enhancement expands the contrast in these local areas so that the background and the objects suppressed by the smoke are now clearly visible. Note that other areas in the image with adequate local contrast are affected to a lesser degree. This is as it should be—we do not want to boost the local contrast (high frequency information) in already "contrasty" areas. Doing so magnifies the scene extrema and therefore reduces the relative global dynamic range available elsewhere.

As an added benefit, the enhanced images have crisper detail because local area contrast enhancement is in fact high frequency emphasis. This high frequency gain is adaptive—being greater in local areas with small local standard deviation.

V. Summary and Discussion

Conventional schemes for local area video contrast stretching suggest complex architectures for real-time implementations which require nonrecursive computation of the local area statistics on a sliding window. The size of the window dictates the hardware complexity, and window sizes between 10 X 10 and 15 X 15 pixels are usually considered optimal. A 15 X 15 window requires simultaneous access to 15 lines of video (14 shift registers or line delays) and 225 adds/pixel at 10-20M pixels/s.

The two-dimensional recursive filter approach in this paper requires just one line delay and one summer (adder) to accomplish the same function. Moreover, the effective size of the local area can be changed simply by changing the filter weights.

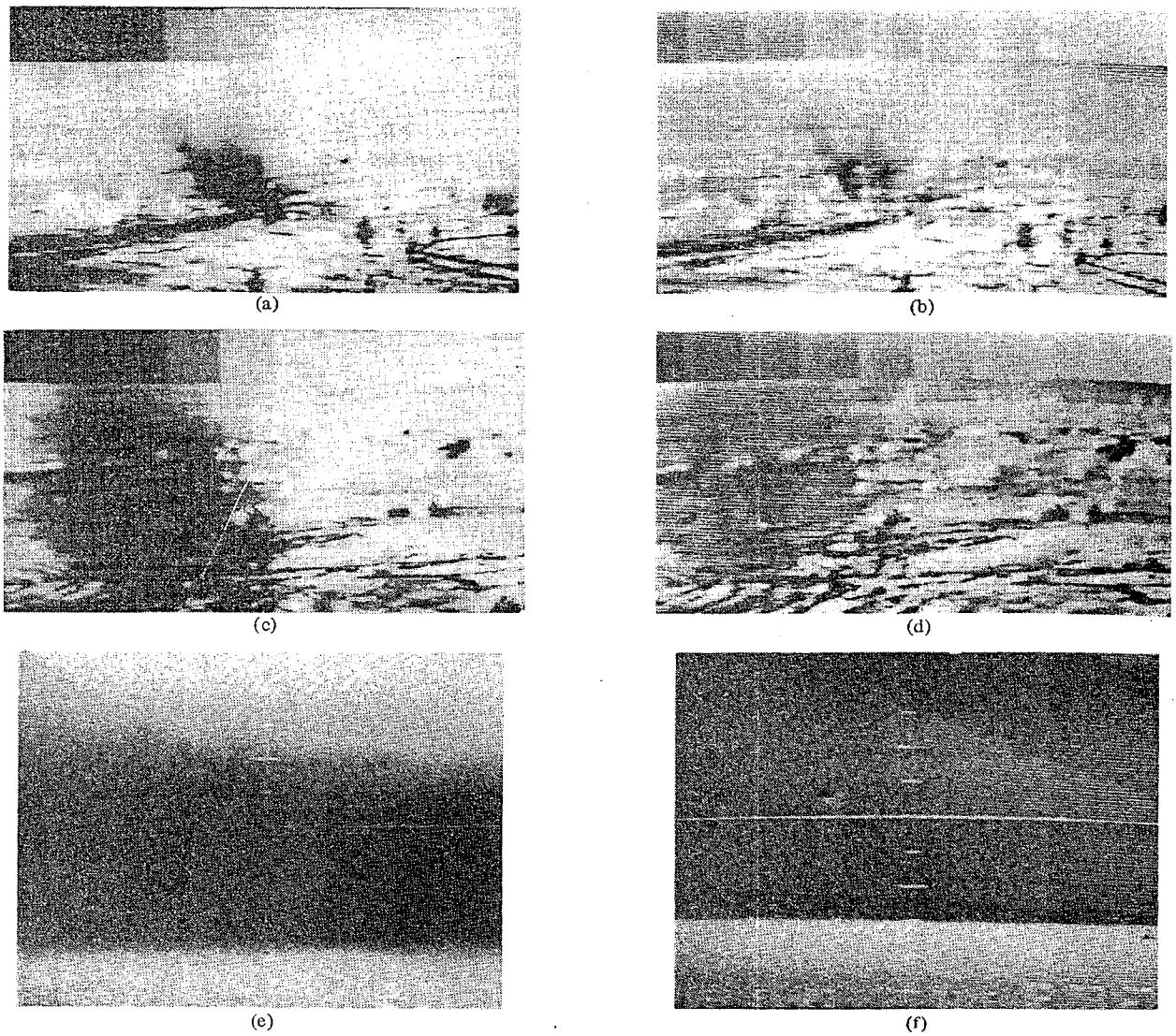


Fig. 8. Results of real-time processing of FLIR video through LACE module.

To change the local are in a nonrecursive sliding window, we would have to change the filter size, i.e., the entire filter structure.

The recursive filter approach would result in much greater simplicity (compared to the nonrecursive approach) in a digital implementation as well. We chose to implement it with sampled analog hardware using CCD line delays, because this approach eliminates high-speed A/D converters, multipliers, and adders at the 10–20 MHz rates encountered in 525 and 875 line video applications. It is interesting to note that the high frequency video does not pass through the CCD line delays in Fig. 6. Therefore, it is possible to use this circuitry without modifications beyond 30 MHz while keeping the CCD clock rates below 15 MHz. This is a boon since commercially available CCD line delays are limited to a 15 MHz clock frequency, but full resolution 875 line video can have bandwidth in excess of 30 MHz. A nonrecursive digital approach would be hard to put to realize these speeds.

This simple hardware occupies only one 6 × 6 in card although built entirely of off-the-shelf components. This image enhancement module has demonstrated the utility of CCD based sampled analog filtering techniques for simple hardware realizations of real-time image enhancement functions.

Acknowledgment

The authors wish to thank J. Dehne of the Night Vision and Electro Optics Laboratory and Dr. M. Geokezas of Honeywell, Inc. for their encouragement and support during this effort. Discussions with Drs. D. H. Tack and J. D. Joseph of Honeywell were very helpful.

References

[1] D. J. Ketcham, "Real-time image enhancement techniques," in *Proc. Soc. Photo-Optical Instrumen. Eng.*, vol. 74 (Image Processing), Feb. 1976, pp. 120–125.

[2] J. L. Harris, "Constant variance enhancement—A digital processing technique," *Appl. Optics*, vol. 16, pp. 1268–1271, May 1977.

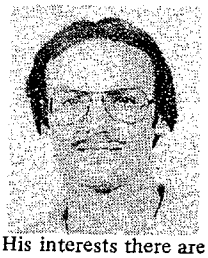
[3] R. C. Gonzalez and B. A. Fittes, "Grey level transformations for interactive image enhancement," *Mechanism and Machine Theory*, vol. 2, pp. 111–122, 1977.

[4] R. E. Woods and R. C. Gonzalez, "Real time digital image enhancement," *Proc. IEEE*, to be published.

Patrenahalli M. Narendra (S'73–M'76) was born in Jog Falls, India, on March 16, 1951. He received the Bachelor of Engineering degree in electronics from Bangalore University in 1971 and the M.S. and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1973 and 1975, respectively.

He has been with Honeywell Systems and Research Center, Minneapolis, MN, since 1976, where he is presently Staff Research Scientist in the Signal and Image Processing Section. His current research interests span dynamic scene analysis, symbolic image processing, and VLSI device architectures for real-time implementation of image processing systems. He is the author of several publications in his areas of interest. In 1979, he won the H. W. Sweatt Award, the highest technical achievement award in Honeywell.

Dr. Narendra is a member of Phi Kappa Phi and the IEEE Computer, PAMI, and ASSP Societies.



Robert C. Fitch (S'73–M'77) was born on June 23, 1953. He received the B.S. degree in electrical engineering from Montana State University, Bozeman, in 1975 and the M.S. degree in electrical engineering from Colorado State University, Fort Collins, in 1977.

He is currently a Senior Research Scientist in the Signal and Image Processing Section at Honeywell Systems and Research Center, Minneapolis, MN, where he has worked since 1977. His interests there are in VLSI implementations for real-time image processing, automatic target recognition, and robotics.

Mr. Fitch is a member of Tau Beta Pi and Phi Kappa Phi.

Determining Surface Orientations of Specular Surfaces by Using the Photometric Stereo Method

KATSUSHI IKEUCHI, MEMBER, IEEE

Abstract—The orientation of patches on the surface of an object can be determined from multiple images taken with different illumination, but from the same viewing position. The method, referred to as photometric stereo, can be implemented using table lookup based on numerical inversion of reflectance maps. Here we concentrate on objects with specularly reflecting surfaces, since these are of importance in industrial applications. Previous methods, intended for diffusely reflecting surfaces, employed point source illumination, which is quite unsuitable in this case. Instead, we use a distributed light source obtained by uneven illumination of a diffusely reflecting planar surface. Experimental results are shown to verify analytic expressions obtained for a method employing three light source distributions.

Index Terms—Bin of bolts and nuts, distributed light source, glossy object, reflectance map, shape from shading, surface inspection.

Manuscript received November 29, 1979; revised February 26, 1981.
The author is with the Computer Vision Section, Electrotechnical Laboratory, Ministry of International Trade and Industry, Ibaraki 305, Japan.

I. Introduction

THIS paper addresses the problem of determining local surface orientation of specular materials from the intensity information under different illumination, but from the same viewing position. This method is referred to as photometric stereo and was first formulated by Woodham [2]. Here, we concentrate on objects with a specularly reflecting surface, since these are of importance in industrial applications. Previous methods [2], [11] intended for diffusely reflecting surfaces employed point source illumination, which is quite unsuitable in this case.

Historically, Horn solved the image intensity equations [11] in order to obtain an object shape from shading information. Horn used a method of characteristic strip expansion for solving the image intensity equation which is a nonlinear first-order partial differential equation. Horn then introduced the reflectance map [3] in order to refine the image intensity

0162-8828/81/1100-0661$00.75

Exhibit I

ISSN 0734-189X

THIS NUMBER COMPLETES VOLUME 24
VOLUME 24, NUMBER 3, DECEMBER 1983

Computer Vision, Graphics, and Image Processing

EDITORS
Linda G. Shapiro
Norman Badler
Herbert Freeman
Thomas S. Huang
Azriel Rosenfeld

AN INTERNATIONAL JOURNAL

ACADEMIC PRESS
A Subsidiary of Harcourt Brace Jovanovich, Publishers
New York London Paris San Diego
San Francisco Sao Paulo Sydney Tokyo Toronto

COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING

Volume 24, Number 3, December 1983

0734-189X/83 $3.00

MADE IN THE UNITED STATES OF AMERICA

Published monthly by Academic Press, Inc., 111 Fifth Avenue, New York, N.Y. 10003.

1983: Volumes 21–24. Price: $260.00 U.S.A. and Canada; $295.00 outside U.S.A. and Canada
1984: Volumes 25–28. Price: $280.00 U.S.A. and Canada; $315.00 outside U.S.A. and Canada
All prices include postage and handling.

Information concerning personal subscription rates may be obtained by writing to the Publisher. All correspondence and subscription orders should be addressed to the office of the Publishers at 111 Fifth Avenue, New York, N.Y. 10003. Send notices of change of address to the office of the Publishers at least 6–8 weeks in advance. Please include both old and new addresses.

Second class postage paid at New York, N.Y., and at additional mailing offices. POSTMASTER: Send changes of address to *Computer Vision, Graphics, and Image Processing*, 111 Fifth Avenue, New York, New York 10003.

COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING

Volume 24, Number 3, December 1983

CONTENTS

COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 24, 363–381 (1983)

SURVEY

Digital Image Enhancement: A Survey*

DAVID C. C. WANG,† ANTHONY H. VAGNUCCI,‡ AND C. C. LI §

†*Bell Telephone Laboratories, Holmdel, New Jersey 07733;* ‡*Department of Medicine, Montefiore Hospital and the University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213;* and §*Department of Electrical Engineering, University of Pittsburgh School of Engineering, Pittsburgh, Pennsylvania 15261*

Received August 24, 1981; revised March 22, 1982 and September 1, 1982

Over decades, many image-enhancement techniques have been proposed. Most of these techniques have been implemented, and their test results have been published. These techniques are surveyed, the underlying concepts briefly described, and their merits discussed. The goal is to help investigators in their selection of enhancement techniques suitable to their needs.

1. INTRODUCTION

The purpose of image enhancement is to improve picture quality. More specifically, it is employed to remove noise, to deblur objects' edges, and to highlight some specified features. Millions of pictures ranging from biomedical images to aerial photographs are generated annually [1]. The application of image enhancement, in general, improves human viewing ability and increases the chance of success in automatic picture processing [2–5]. An example is shown in Fig. 1-1 which is a computerized tomogram of a cross section of the head of a patient with a pituitary tumor (indicated by the arrow) [6]. In Fig. 1-1(b), the original picture (Fig. 1-1(a)) is degraded by additive noise such that the tumor is not visible. After the application of a smoothing scheme, the tumor becomes again clear as shown in Figs. 1-1(c) and (d).

Over decades, many image-enhancement techniques have been proposed. Most of these techniques have been implemented, and their test results have been published. In this paper, we survey these techniques, briefly describe the underlying concepts, and discuss their merits. The goal of this survey is to help investigators in their selection of enhancement techniques suitable to their needs.

Image restoration is considered as image enhancement by some authors; it is not, however, included in this paper. Image restoration attempts to estimate the "true" image of an observed picture given, if possible, a priori knowledge of the degradation [7]. Its emphasis is on degradation modeling and on image recovery by inversion of the degrading process [8]. Its goal is to obtain a picture with "enhanced" objects which differ as little as possible from those in the "true" image. Optimization of some prespecified objective criteria is thus required. On the other hand, image enhancement, to be discussed in this paper, is designed to manipulate the image on

*Supported in part by a grant from the Western Pennsylvania Heart Association. Work done while D. C. C. Wang was with the University of Pittsburgh School of Medicine.

0734-189X/83 $3.00

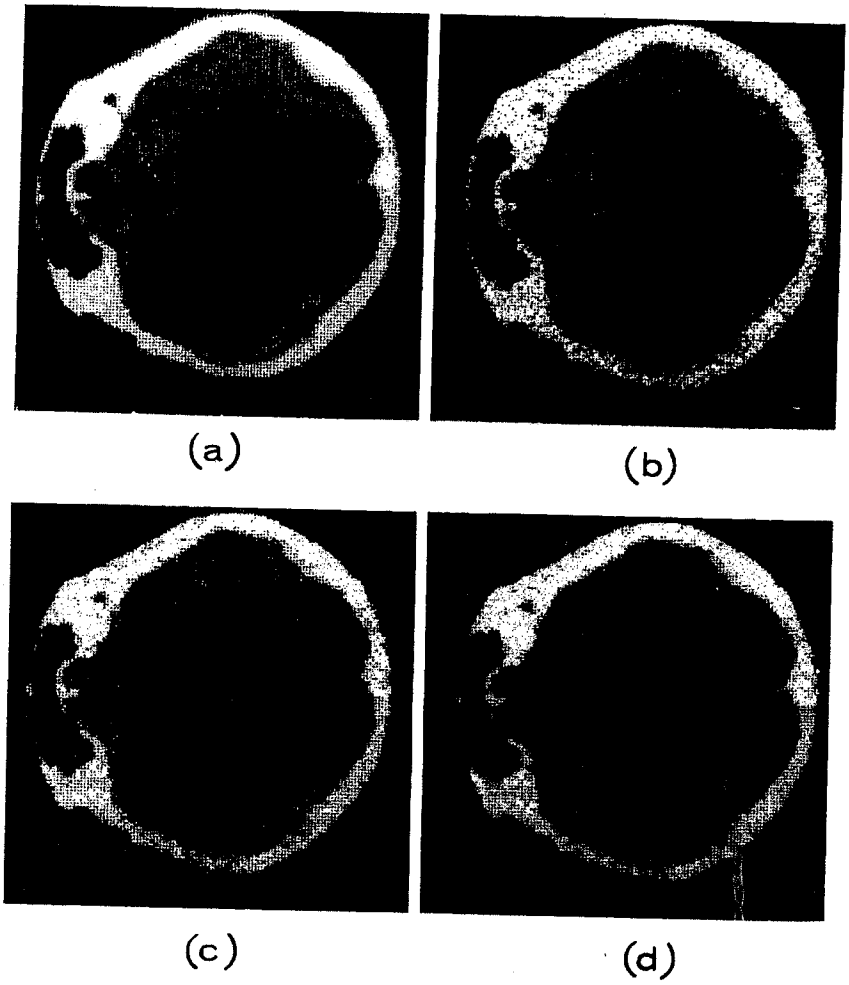


FIG. 1-1. (a) CT scan of a cross section of the head of a patient with a pituitary tumor (indicated by arrow). (b) Image degraded by additive noise. (c) and (d) The degraded image smoothed three and five times, respectively. See Wang *et al.* [6]. Reproduced by permission of the publisher.

the basis of the psychophysical characteristics of the human visual system [8]. It may even "distort" the image by deemphasis of irrelevant materials and enhancement of features of interest [9, 10].

Most of the existing image-enhancement techniques are heuristic and problem oriented [11]. Each technique is devised to handle some special kind of pictures. Modeling of the degradation process, in general, is not required; however, knowledge of the degradation phenomena may give a guide to investigators in their choice of enhancement techniques. For monochromatic images, the degradation may be categorized into point degradation, spatial degradation, and their combination. Point degradations, which include additive and multiplicative noise, do not blur the

image, but distort the gray levels of pixels. These degradations might be due to film grain, lens shading, scatter light, and other optical sensor defects. Spatial degradations, on the other hand, smear the image so that it loses resolution, and the transition from object to background becomes unclear. These degradations might be due to diffraction, optical system aberrations, defocusing, object motion, etc. For a more detailed discussion of sources of degradations, the reader is referred to some review articles and books [12–17].

Since most of the image-enhancement techniques are problem oriented, the success of the application of a technique to a particular image depends on the subjective judgment of the viewer [11]. No universal quantitative evaluation criterion is available currently. Readers are advised to apply several appropriate techniques to their images, and to choose the best one, based on their own judgment, for that image which best fits their purpose. In this paper, the enhancement techniques are evaluated on the basis of their practical applicability and computational complexity. Because computation facilities vary, some techniques may be considered too complex by some investigators, not so by others.

In Section 2, we define the notation used in this paper. In Section 3, we outline a classification of the existing techniques according to their properties. We then survey these techniques, according to the suggested classification, in Sections 4 to 7.

2. NOTATION

The notation used in this paper is listed below:

$g(x, y)$	Gray level of the pixel at coordinates (x, y) of the observed image;
$g'(x, y)$	Gray level of the pixel at (x, y) of the enhanced image;
$\mathbf{W}$	A 3×3 weighting-factor matrix for smoothing where each element is a weighting factor;
$P(g(x, y))$	Value of cumulative distribution function of an observed image at gray level $g(x, y)$;
$\nabla^2 g(x, y)$	Laplacian of the observed image at (x, y);
$g_{\max}$	Maximum gray level of the observed image;
$g_{\min}$	Minimum gray level of the observed image;
$G(f_x, f_y)$	Fourier-transformed image of the observed image;
$G'(f_x, f_y)$	Fourier-transformed image of the enhanced image;
$H(f_x, f_y)$	Frequency domain filtering function.

3. CLASSIFICATION OF ENHANCEMENT TECHNIQUES

Image enhancement is achieved by image processing through operators. The existing techniques can be categorized on the basis of the properties of their operators. Classifications are based on (1) the operator's sensitivity to image context, (2) the area which the operation covers, (3) the goals of the operation, and (4) the technical methods involved.

According to the operator's sensitivity to image context, enhancement techniques can be classified as (a) context-free, and (b) context-sensitive [4]. A *context-free*

technique provides a position-invariant operator in which all the parameters are fixed a priori. A *context-sensitive* technique furnishes a position-variant operator in which the parameters change in response to local image characteristics. A context-free operator is computationally simpler to apply. However, for pictures with variable information content, a context-sensitive operator is more suitable.

In accordance with the area covered by the operator, the existing techniques can be divided into local and global. *Local operators* take a subimage into consideration at each time; they can be further subdivided into fixed-sized and variable-sized. In a *global operation*, the whole image is involved. Computationally, application of a local operator requires less storage than that of a global operator.

Based on their goals, the existing techniques can be grouped into (a) noise cleaning, (b) feature enhancement, and (c) noise cleaning plus feature enhancement. It is well known that a picture can be degraded by additive noise, and be smeared by point-spread functions. *The noise-cleaning operator* aims at the removal of random noise; in other words, it disregards irrelevant information. The *feature-enhancement operator* attempts to decrease the blurring, and to reveal the features of interest. These two operators deal with different degradation phenomena; in practice, however, many operators are a combination of both.

According to the technical methods involved, the published techniques can be organized into four groups. They are

(a) *Spatial smoothing* of regions, which employs linear or nonlinear spatial-domain low-pass filters;

(b) *Gray-level rescaling*, which manipulates or requantizes gray levels for contrast enhancement;

(c) *Edge-enhancement*, which involves linear or nonlinear spatial-domain high-pass filters;

(d) *Frequency-domain filtering*, which utilizes low- or high-pass filters in the frequency domain where Fourier transformation is required.

The last classification is more comprehensive than the other three listed above. It specifies the technical tools which implicitly indicate the goal of the technique. In the following sections, we shall survey the existing techniques according to this classification.

4. SPATIAL SMOOTHING OF REGIONS

In this section, the observed image is assumed to be the true image degraded by random noise. Although it may not be explicitly stated, most authors who employ spatial smoothing implicitly assume that the random noise is additive and normally distributed with zero mean. Thus, smoothing, linear or nonlinear, over a region tends to remove the noise at each pixel.

The techniques described in this section provide local operators, unless otherwise stated. The required storage is usually small. The goal is noise cleaning. Some of the techniques even smear the observed image. Thus, edge enhancement may be needed afterward.

The simplest smoothing technique is equal-weighted averaging over a neighborhood [9, 10, 18, 19]. It can be expressed as follows:

$$g'(x, y) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} w(i, j)\ g(x - i, y - j) \tag{4-1}$$

where the weighting factors $w(i, j)$ are equal and

$$w(i, j) = \frac{1}{(2m + 1)(2n + 1)}. \tag{4-2}$$

In (4-1), the gray level at (x, y) is replaced by the gray level average over a $(2m + 1)$ by $(2n + 1)$ rectangular neighborhood surrounding (x, y). This simple smoothing scheme can remove noise efficiently; it will, however, blur the picture, especially at the edges of objects. Blurring becomes more severe as m and n increase.

To reduce the blurring effect, several unequal-weighted smoothing techniques have been introduced. These techniques weight the gray level at (x, y) more than those of its neighbors. Graham [20] used a 3×3 neighborhood, and the weighting factor matrix

$$\mathbf{W} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix} \tag{4-3}$$

Brown [21] proposed the weighting factor matrix

$$\mathbf{W} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \tag{4-4}$$

In formulas (4-2) to (4-4), the weighting-factor matrices are context-free. They are easy to implement. However, (4-2) tends to smear the sharpness of the edge severely. Procedures (4-3) and (4-4) may not remove the noise as efficiently as (4-2); they do blur the edge, although not to the same extent as (4-2).

To improve the results obtained by the above schemes, Lev *et al.* [22] proposed a context-sensitive weighting-factor matrix. Let $g(x, y) = e$, and the gray levels of its neighbors be

$$\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}$$

The proposed weighting-factor matrix is

$$\mathbf{W} = \tfrac{1}{9} \begin{bmatrix} \alpha\gamma\varepsilon & \alpha\gamma\eta & \alpha\xi\eta \\ \alpha\varepsilon\theta & 1 & \delta\xi\eta \\ \beta\varepsilon\theta & \beta\delta\theta & \beta\delta\xi \end{bmatrix} \tag{4-5}$$

where

$$\alpha = \exp(-|(a+b+c)-(d+e+f)|/\sigma),$$
$$\beta = \exp(-|(g+h+i)-(d+e+f)|/\sigma),$$
$$\gamma = \exp(-|(a+b+d)-(c+e+g)|/\sigma),$$
etc.

and σ is a prespecified parameter. In (4-5), the weighting is reduced in the direction which shows a sharp change in the gray level. The homogeneous area is weighted more. The blurring effect at the edge can be decreased. This technique does yield better results. Computationally, it is more complicated than the previous three. It also requires the specification of σ, which may need a priori knowledge or some trials to obtain the best result.

According to criteria similar to those suggested by Lev *et al.* [22], Wang *et al.* [6] proposed a computationally simpler 3×3 weighting-factor matrix. The matrix is

$$\mathbf{W} = \begin{bmatrix} w(x-1,y-1) & w(x-1,y) & w(x-1,y+1) \\ w(x,y-1) & w(x,y) & w(x,y+1) \\ w(x+1,y-1) & w(x+1,y) & w(x+1,y+1) \end{bmatrix} \tag{4-6}$$

where

$$w(x+k,y+l) = \tfrac{1}{2}\left[\sum_k \sum_l \delta(x,y;k,l)\right]^{-1} \delta(x,y;k,l) \qquad l,k = -1,0,1$$
$$w(x,y) = \tfrac{1}{2}$$

and

$$\delta(x,y;k,l) = 1/|g(x+k,y+l) - g(x,y)|$$

This technique smoothes images with very little blurring effect. It requires no prespecified parameters. However, it may not remove the noise as rapidly as the first three techniques presented above.

Graham [12] defined two measures of flexure Δ_x and Δ_y, which are approximations to the second-order derivatives of pixel gray level in the x and y directions, respectively. He then replaced $g(x,y)$ by $g'(x,y)$ where

$$g'(x,y) = \begin{cases} \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} g(x-i,y-j), & \text{if } \Delta_x \le T \text{ and } \Delta_y \le T \\ \frac{1}{3}\sum_{j=-1}^{1} g(x,y-j), & \text{if } \Delta_x > T \text{ and } \Delta_y \le T \\ \frac{1}{3}\sum_{i=-1}^{1} g(x-i,y), & \text{if } \Delta_x \le T \text{ and } \Delta_y \ge T \\ g(x,y), & \text{if } \Delta_x \ge T \text{ and } \Delta_y > T \end{cases} \tag{4-7}$$

and T is a prespecified threshold parameter. This scheme smooths only the homoge-

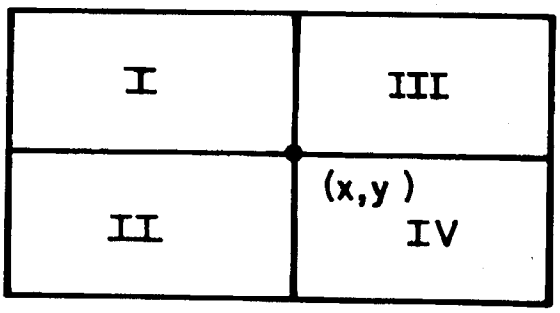


FIG. 4-1. Four neighboring regions used by Kuwahara *et al.* [23].

neous area, that is, the area with a small second derivative. It does not smooth the edge pixels. However, it ignores those isolated noise pixels which are much darker or brighter than their neighbors. In addition, this technique depends on the specification of the parameter T which requires a priori knowledge or some trials.

Another smoothing scheme suitable to clean up "pepper and salt" noise, that is, isolated noise, has been suggested [9, 15, 18, 20]. In this scheme, the gray levels of those pixels which are much darker or much brighter than their neighbors' average are replaced by the average gray level. Thus, only isolated dark points in a light region and isolated light points in a dark region are smoothed. This scheme also requires the specification of the threshold.

Kuwahara *et al.* [23] and Tomita *et al.* [24] proposed a smoothing scheme which replaces the gray level at (x, y) by the average gray level of its most homogeneous neighboring region. Kuwahara *et al.* [23] divided the neighborhood into four regions as shown in Fig. (4-1), and Tomita *et al.* [24] partitioned it into five regions, as illustrated in Fig. (4-2). They then calculated the average and variance of gray levels within each region. The region which has the smallest variance is considered as the region which does not include boundaries. The gray level at (x, y) is then replaced by the average of this region. This technique can achieve both noise removal and edge enhancement since the smoothing operation does preserve the image's edges. This scheme does not, however, yield good results when applied to images with

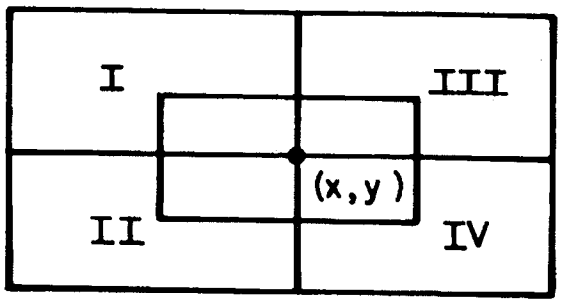


FIG. 4-2. Five neighboring regions employed by Tomita *et al.* [24].

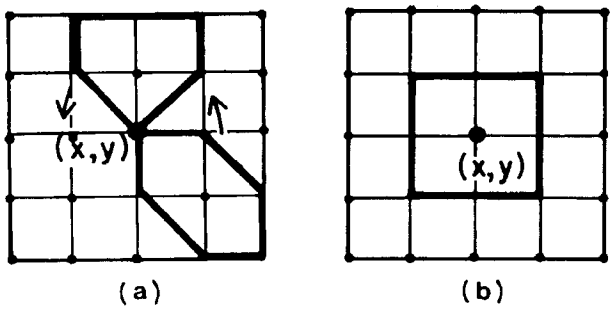


FIG. 4-3. (a) Four pentagonal and four hexagonal regions. (b) A square region used in [25].

complex-shaped regions because it employs four or more large rectangular areas as neighborhood regions.

A revision and improvement of the last two techniques was published by Nagao and Matsuyama [25]. These authors divide the 5×5 neighborhood of (x, y) into nine regions as shown in Fig. (4-3). The nine regions are four pentagonal and hexagonal regions (Fig. 4-3(a)), and a 3×3 square region (Fig. 4-3(b)). By the same criteria used in the previous techniques [23, 24], $g(x, y)$ is replaced by the average of the region which has the smallest variance. This smoothing scheme is iterative; the iteration stops when no change of the gray level occurs. The results thus obtained present a significant improvement over those derived from the previously mentioned techniques [23, 24]. However, the computation of this scheme is far more complex and time consuming; also it may smooth out the fine features of the image.

Anderson and Netravali [26] suggested a weighted smoothing procedure based on subjective criteria. The weighting factors are found by optimizing an objective function which maximizes noise suppression and minimizes blurring. This scheme has been simplified by Trussell [27]. Since human eyes can tolerate more noise in areas of high signal strength, Trussell's scheme leaves the image alone in areas of high activity, and smooths the noisy image in areas of low activity. Let the difference between the observed image $g(x, y)$ and the blurred (smoothed) image $b(x, y)$ be D. If the smoothed image $b(x, y)$ perfectly fits the "true" picture, then $\sigma_D^2 = \sigma_n^2$. In other words, the variance of D equals the variance of the noise. If $\sigma_D^2 \leq \sigma_n^2$, the smoothed image is an adequate representation of the "true" image. Otherwise, it is not adequate. Trussell interactively selects a low activity area, and calculates the variance of that area and lets the variance be σ_n^2. For each pixel (x, y) and its neighbors, he calculates $\sigma_D^2(x, y)$. Then $g(x, y)$ will be replaced by

$$g'(x, y) = \theta(x, y)b(x, y) + (1 - \theta(x - y))g(x, y) \tag{4-8}$$

where

$$\theta(x, y) = \min\{1.0, \sigma_n^2/\sigma_D^2(x, y)\}.$$

Those areas which have higher activities, i.e., large σ_D^2, show less deviation from the observed image. This technique is suitable for many images; however, its effectiveness is weakened if the noise is not uniformly distributed across the whole image.

A nonlinear smoothing scheme, called median filtering, was first proposed by Tukey [28], and then adapted by others [15, 29–32]. The filtering operation replaces $g(x, y)$ by the median gray level of a $(2m + 1)$ by $(2n + 1)$ neighborhood surrounding (x, y). This smoothing technique performs satisfactorily in step functions and ramp functions. However, the pulse functions whose periods are less than n (i.e., one-half the window width) are suppressed. Thus, in some cases, the median filter provides noise removal, and in other cases, it may cause signal suppression. To remedy this, Pratt [15] suggests using variable window widths. He starts the filtering operation by letting $n = 1$, and repeats the operation by increasing n. The process is terminated when the filtering begins to do more harm than good. This revised median filtering may improve the result; however, it is time consuming, and different n's for different regions in one image may be needed.

To avoid the noise introduced during the digitization of the picture, averaging over multiple digitized copies has been suggested [9, 18, 33]. The arithmetic mean of each pixel over the copies is employed to eliminate additive noise which is due to light scattering, electronic scattering, etc. The geometric mean is utilized to clean multiplicative noise which is due to the randomness of film grain formation, lens defects, etc.

5. GRAY LEVEL RESCALING

Gray level rescaling directly requantizes or maps each pixel to a new gray level to improve the contrast of a picture. Image-enhancement techniques which utilize gray level rescaling are different from those techniques presented in the last section. In general, gray level rescaling takes into consideration one pixel at a time, and is independent of its neighbors. The goal is object enhancement, and most of the operations are global.

In many pictures, the gray levels of the objects are so close to that of the background that it is difficult to discriminate them. Contrast enhancement is thus needed to increase the gray level differences between objects and background. In other situations, the gray levels of a large percentage of pixels concentrate in a narrow portion of the histogram. This makes the fine details hardly visible. Requantization or rescaling to increase dynamic range and to single out hidden fine structures is thus required.

The simplest technique of this category is a functional mapping of the gray level $g(x, y)$ [3, 9, 12, 34–36], i.e.,

$$g'(x, y) = f(g(x, y)) \tag{5-1}$$

where $f(\cdot)$ is a prespecified function. Some of the most popular mapping functions are illustrated in Fig. 5-1. Figures 5-1a to d are piecewise linear functions, and Fig. 5-1e is nonlinear. In Fig. 5-1a, the dark area is stretched, and the bright area is compressed. Thus, the contrast in the dark area is increased. In 5-1b, the operator performs opposite to that of 5-1a, and in 5-1c it stretches the midrange gray levels. The operator of Fig. 5-1d carries out a level slicing in which the two most significant bits are removed. Figure 5-1e shows two gamma curve corrections for film and display nonlinearities. These mappings are very easy to implement, and usually yield satisfying results. However, trial and error is needed to obtain the best results. In many cases, the investigator has to provide his own mapping functions.

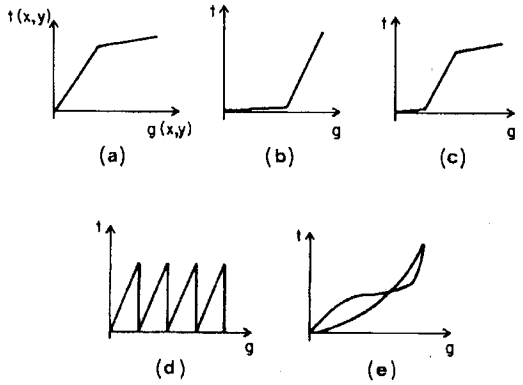


FIG. 5-1. Gray level rescaling functions: (a) dark area stretched; (b) light area stretched; (c) midrange stretched; (d) level slicing; (e) gamma function.

The gray level histograms of most images, in general, show bright or dark peaks. Figure 5-2a exhibits a histogram skewed toward dark gray levels; the peak represents a large dark area which may contain some information with fine details. To enhance this information, a technique called histogram linearization has been proposed and adopted [15, 37–43]. This technique tends to map the observed gray levels into new gray levels such that the new picture has a uniform gray level histogram. The operation is equivalent to maximizing the zeroth-order entropy. It has been suggested [44] that the information in the observed image is thus maximized for the observer; in the new picture, the dynamic range of the dark area will be expanded. Curve 1 of Fig. 5-2b is the cumulative distribution function (CDF) of a picture with

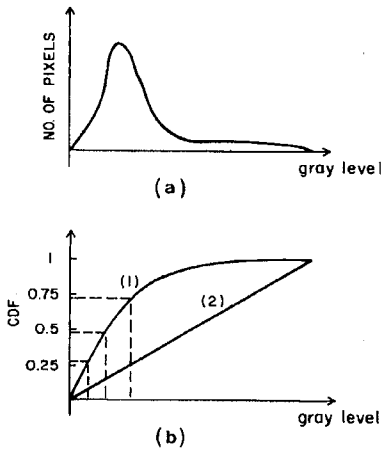


FIG. 5-2. (a) Gray level histogram of a typical image. (b) Curve 1 is the cumulative distribution function (CDF) of (a). Curve 2 is the desired linear CDF. See text.

continuous gray levels. Curve 2 is the desired linear CDF. The procedure to linearize the observed CDF (curve 1) consists of the following steps: (a) selection of equal intervals in the ordinates as shown in Fig. 5-2b; (b) from curve 1, map the intervals in the ordinate to the gray levels in the abscissa; (c) requantize the observed picture. This procedure can be formulated as [15]

$$g'(x, y) = [g_{\max} - g_{\min}]P(g(x, y)) + g_{\min} \tag{5-2}$$

where $P(g(x, y))$ is the CDF, and $g_{\max}$, $g_{\min}$ are the maximum and minimum gray levels in the observed image, respectively.

In the above procedure, an exact linearization is a practical impossibility because the gray levels are discrete; an approximate linearization of the CDF would yield a smaller number of gray levels in the enhanced image. This would create a larger quantization error [15]. A technique which performs histogram linearization with the same number of gray levels for the observed and for the enhanced images is required.

Suppose that there are n pixels in the observed image with gray level l (i.e., bin l in the histogram). Suppose also that in order to create a uniform histogram without reduction of the gray level numbers, this bin has to be broken up such that n' pixels are assigned to gray level l', and the remaining pixels to level $l' + 1$. A simple way to do so is to choose these n' pixels randomly from the n pixels. To avoid randomness in the selection of the pixels, a procedure called GRATRN has been proposed [38]. In this procedure, the average gray level s of the neighborhood around each of these n pixels is calculated. A CDF of these n average gray levels is constructed. Select, then, an average gray level s', such that the number of pixels with $s \leq s'$ equals n'. These n' pixels are then assigned to a new gray level l'. The remaining pixels with $s \geq s'$ are assigned to $l' + 1$. GRATN breaks each bin according to the average of the neighboring pixels of each pixel in that bin; the breaking, however, is somehow arbitrary, particularly for those pixels at the edge. This technique is time consuming, especially in those cases where the number of bins is very large.

It has been pointed out that human perception performs a nonlinear transformation of the light intensity [44–47]. To include human perception in the image enhancement, histogram hyperbolization [15, 44] is preferred to histogram linearization. The desired histogram of the computer enhanced image is thus of the hyperbolic form. The transforms are as follows [15]:

$$\begin{aligned}
&g'(x, y) = g_{\min} + \frac{1}{\alpha}\ln[1 - P(g(x, y))] \quad \text{for exponential histogram} \\
&g'(x, y) = \left[\left(g_{\max}^{1/3} - g_{\min}^{1/3}\right)P(g(x, y)) + g_{\min}^{1/3}\right]^3 \quad \text{for cubic root output histogram} \\
&g'(x, y) = g_{\min}\left[\frac{g_{\max}}{g_{\min}}\right]^{P(g(x, y))} \quad \text{for logarithmic output histogram}
\end{aligned} \tag{5-3}$$

where α is a constant. According to the perception performed by the rods and cones of the human retina, the perceived image would have a uniform histogram.

A scheme called iterative histogram thinning, which achieves results opposite to those obtained by histogram linearization, is proposed to thin each peak in the histogram [48, 49]. Let the number of pixels in bin i of the histogram be B_i. This scheme examines each bin i and its neighboring 2γ bins on two sides. If B_i is greater than the average A of $B_{i+1},\ldots,B_{i+j}$, a ratio $\beta = (B_i - A)/B_i$ is computed. Then the gray levels of $B_{i+j}\beta$ pixels are changed from $i + j$ to $i + j - 1, j = 1,2, \ldots,\gamma$. Thus, B_i increases and its neighboring bins decrease. This technique yields images with homogeneous objects and background with enhanced contrast; it may, however, result in quantization error.

A noise-cheating image-enhancement scheme has been proposed [50, 51] to suppress noise and to enhance contrast. A 4×4 average (not moving average) over the observed image is performed, and followed by a 2×2 average. A 512×512 image is thus reduced to 64×64. The averaged image is then requantized such that the difference between two consecutive gray levels is at least four standard deviations. In the requantized image, isolated points are deleted. The 4×4 averaged image is then quantized. Each pixel in the 4×4 averaged image is assigned a value chosen from its surrounding points in the quantized 8×8 averaged image. The value is chosen such that the difference between before and after quantization is a minimum. Good results have been obtained by this technique; it is, however, a complex procedure, and extremely small objects may be lost in the averaging step.

Image subtraction is commonly used in serial angiography [52–55] to deemphasize the irrelevant material and to enhance the material of interest. The original radiograph is subtracted from the subsequent image. The subsequent image is obtained in the same way as the original, except that a radio-opaque substance is injected into an artery of the patient. Thus, the blood vessels are enhanced while bony structures are unchanged. Hence, image subtraction can eliminate bones, and enhance the blood vessels. The earlier work was done optically. Hall, *et al.* [39] discussed image subtraction by digital procedure. Recent developments are reviewed in [78].

6. EDGE ENHANCEMENT

In edge enhancement, we attempt to deblur the edge of an object within an image. In other words, the technique is employed to increase the gray level difference between the edge pixels of the object and of its neighboring background. Most of the edge-enhancement operators are context-free and local.

Effects which blur pictures are diffusion, defocusing, and object motion. The gray level profile at an "ideal" edge is a step function. Diffusion and defocusing blur the edge into a ramp function. The gray level at the edge of an object is decreased, and that at its neighboring background is increased. The contrast at the edge is thus reduced or even lost. To eliminate these blurring effects, an antidiffusion operator has been proposed and adopted [9, 15, 56–61]. To eliminate the effect of object motion, image restoration can be applied [14].

An anti-diffusion operation makes up for the loss of gray level at the object edge and subtracts the increment of gray level at its neighboring background. Figures 6-1a and b show a perfect one-dimensional step function $u(x)$ and its second-order derivative $\nabla^2 u(x)$, respectively. By subtracting $\nabla^2 u(x)$ from $u(x)$, we obtain the graph illustrated in Fig. 6-1c. It can be seen that the gray level difference between the edge ($x = 0^+$) and its neighboring background ($x = 0^-$) is increased. The overshooting of gray levels at the edge will enhance the edge. The anti-diffusion

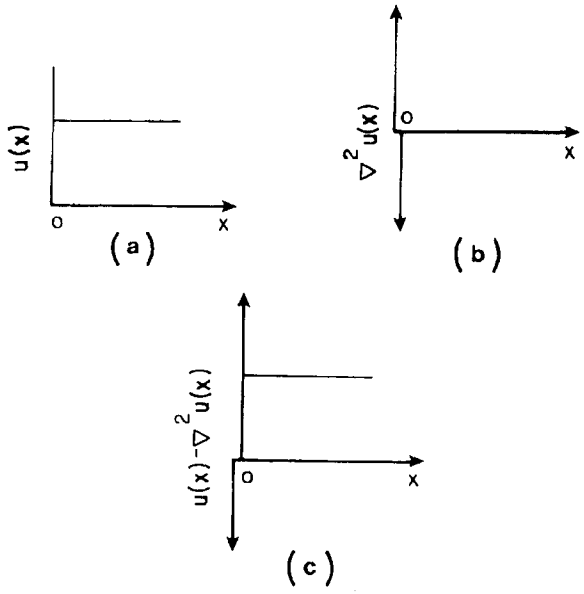


FIG. 6-1. (a) A step function $u(x)$; (b) the second derivative of $u(x)$, $\nabla^2 u(x)$; (c) $u(x) - \nabla^2 u(x)$.

operation can be formulated as

$$g'(x, y) = g(x, y) - \gamma \nabla^2 g(x, y) \tag{6-1}$$

where γ is a constant. Figure 6-1 illustrates a 1-dim. example of edge enhancement by an anti-diffusion operation on a smeared edge. Digital implementation of Eq. (6-1) is a matrix which has both positive and negative elements [9, 15, 57, 58]. This technique may yield a clear view of the edges; it may also enhance the noise.

Overemphasis on accidental fluctuations is one of the major problems with edge-enhancement techniques. A possible solution can be offered by edge enhancement along the direction of the local gradient [56]. This technique enhances the edge according to

$$g'(x, y) = g(x, y) - \gamma \frac{\delta^2 g(x, y)}{\delta n^2} \tag{6-2}$$

where

$$\frac{\delta^2 g(x, y)}{\delta n^2} = \frac{\frac{\delta^2 g}{\delta x^2}\left(\frac{\delta g}{\delta x}\right)^2 + 2\frac{\delta^2 g}{\delta x \delta y}\frac{\delta g}{\delta x}\frac{\delta g}{\delta x}\frac{\delta^2 g}{\delta y^2}\left(\frac{\delta g}{\delta y}\right)^2}{\left(\frac{\delta g}{\delta x}\right)^2 + \left(\frac{\delta g}{\delta y}\right)^2}.$$

It should be noted that the small scale curvature is probably due to noise. A modification of the above technique which enhances the contour along the gradient, and smooths the edge along the contour, has been proposed [62]. It is formulated as

follows:

$$g'(x, y) = g(x, y) - \gamma\left(\frac{\delta^2 g(x, y)}{\delta n^2} - \frac{1}{3}\frac{\delta g(x, y)}{\delta t^2}\right)$$

where

$$\frac{\delta^2 g(x, y)}{\delta t^2} = \frac{\frac{\delta^2 g}{\delta x^2}\left(\frac{\delta g}{\delta y}\right)^2 - 2\frac{\delta^2 g}{\delta x \delta y}\frac{\delta g}{\delta x}\frac{\delta g}{\delta y} + \frac{\delta^2 g}{\delta y^2}\left(\frac{\delta g}{\delta x}\right)^2}{\left(\frac{\delta g}{\delta x}\right)^2 + \left(\frac{\delta g}{\delta y}\right)^2} \tag{6-3}$$

which is the second-order derivative along the tangent of the contour. The computation is complicated. However, it does enhance the slope and suppress the accidental jaggedness.

A filtering technique using local statistics was first proposed by Wallis [63], and then extended by Lee [64, 65]. This technique tends to enhance subtle details. The algorithm is

$$g'(x, y) = \bar{g}(x, y) + k(g(x, y) - \bar{g}(x, y)) \tag{6-4}$$

where $\bar{g}(x, y)$ is the local gray level mean surrounding the pixel (x, y). It should be noted that in Eq. (6-4), if $k > 1$, the difference between the local mean and gray level at (x, y) is magnified, and the edge will be sharpened as if passed through a high-pass filter. When $k = 2$, Eq. (6-4) is equivalent to Eq. (6-1). If $0 < k < 1$, the difference is diminished, and the image will be smoothed as if acted upon by a low-pass filter. If $k = 0$, the operation is just a simple smoothing. Experiments [64] show that when k is large, all the fine details, including noise, are enhanced.

Another technique which emphasizes edges and cleans "salt and pepper" noise has been presented [43]. At the isolated noise pixels, the Laplacian $|\nabla^2 g(x, y)|$'s are very large. This technique performs an anti-diffusion operation as in Eq. (6-1) if $|\nabla^2 g(x, y)|$ is less than a prespecified threshold T. It smooths a pixel as Eq. (4-1) if $|\nabla^2 g(x, y)|$ is larger than the threshold. In other words,

$$g'(x, y) = \begin{cases} g(x, y) - \gamma\nabla^2 g(x, y) & \text{if } |\nabla^2 g(x, y)| \le T \\ \dfrac{1}{(2m+1)(2n+1)} \displaystyle\sum_{i=-m}^{m} \sum_{j=-n}^{m} g(x-i, y-j) & \text{if } |\nabla^2 g(x, y)| > T. \end{cases} \tag{6-5}$$

As shown in this equation, the enhancement technique requires the specification of γ and of threshold T; thus many trials are needed.

An anti-diffusion operation which takes advantage of electronic scanning has been suggested [66–69]. The basic idea is to simultaneously scan with two scanning spots, one fine and one coarse, and subtract the resulting images electronically. Let the fine

resolution image be $g(x, y)$, and the coarse resolution image be $\bar{g}(x, y)$. The algorithm of this technique can be described by Eq. (6-4).

7. FILTERING IN THE FREQUENCY DOMAIN

An image can be viewed as a matrix of points with each point represented by some gray level. It can also be considered as a combination of two-dimensional Fourier series with different frequencies. Thus, a picture can be represented uniquely in the frequency domain. Transformation from spatial domain to frequency domain can be achieved by the fast Fourier transform (FFT) algorithm [70–72].

It has been noticed [73] that low-frequency Fourier components correspond to a homogeneous object and background, and high frequencies are introduced by the occurrence of a sharp edge or noise. Thus, low-pass or high-pass filters can be employed to smooth the picture, or to enhance the sharp edges. The design of these filters has been well developed in digital signal processing [74, 75]. Figure 7.1 illustrates a block diagram of a filter in the frequency domain. The operation is global and context-free.

Duda and Hart [73] considered as a low-pass filter

$$H(f_x, f_y) = \left[(\cos \pi f_x)(\cos \pi f_y)\right]^{\alpha}, \quad \alpha > 1 \tag{7-1}$$

and as a high-pass filter

$$H(f_x, f_y) = 1 - \left[(\cos \pi f_x)(\cos \pi f_y)\right]^{\alpha} \quad \alpha \geq 1. \tag{7-2}$$

It has been shown that a low-pass filter removes the noise; however, it blurs the picture. High-pass filtering sharpens the edges, but it enhances the noise and roughens the object. A compromise filter, called a high-frequency-emphasis (HFE) filter, is proposed. This filter keeps the low frequency unchanged, and exaggerates the high frequency. Duda and Hart proposed the following HFE filter:

$$H(f_x, f_y) = 2.0 - \left[(\cos \pi f_x)(\cos \pi f_y)\right]^{\alpha} \quad \alpha \geq 1. \tag{7-3}$$

The Butterworth filter [74] with a cutoff frequency f_0 has been utilized by Gonzalez and Wintz [8]. Butterworth proposed a low-pass filter

$$H(f_x, f_y) = \frac{1}{1 + \left[\left(f_x^2 + f_x^2\right)^{1/2} / f_0\right]^{2n}}, \tag{7-4}$$

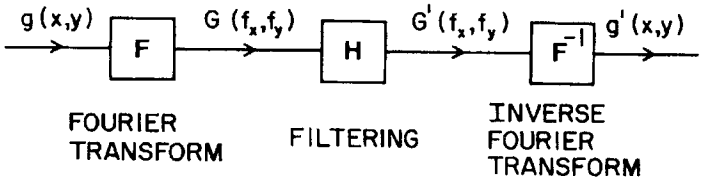


FIG. 7-1. A block diagram of filtering in the frequency domain.

a high-pass filter

$$H(f_x, f_y) = \frac{1}{1 + \left[f_0 / \left(f_x^2 + f_y^2\right)^{1/2}\right]^{2n}}, \tag{7-5}$$

and a high-frequency-emphasis filter

$$H(f_x, f_y) = 1 + \frac{1}{1 + \left[f_0 / \left(f_x^2 + f_y^2\right)^{1/2}\right]^{2n}}. \tag{7-6}$$

Hall *et al.* [39] applied an exponential filter to enhance radiographic images. Their exponential low-pass filter is given by

$$H(f_x, f_y) = \exp\left[-\left(f_x^2 + f_y^2\right)^{1/2} / f_0\right]^n, \tag{7-7}$$

their high-pass filter is represented by

$$H(f_x, f_y) = \exp\left[-f_0 / \left(f_x^2 + f_y^2\right)^{1/2}\right]^n, \tag{7-8}$$

and their high-frequency-emphasis filter is

$$H(f_x, f_y) = 1 + \exp\left[-f_0 / \left(f_x^2 + f_y^2\right)^{1/2}\right]^n. \tag{7-9}$$

Most of the energy of a picture is concentrated in the low-frequency portion. Hence, a logarithmic filter in the frequency domain is proposed [37]. Let

$$G(f_x, f_y) = |G(f_x, f_y)| e^{j\Phi(f_x, f_y)} \tag{7-10}$$

which, after filtering, becomes

$$G'(f_x, f_y) = \left(\log|G(f_x, f_y)|\right) e^{j\Phi(f_x, f_y)}. \tag{7-11}$$

Thus, the low-frequency portion is reduced more significantly than the high-frequency portion; high frequencies are emphasized more than low frequencies.

A similar filter, called α processing, has been suggested by Andrews [76]. The operation is

$$G'(f_x, f_y) = |G(f_x, f_y)|^{\alpha} e^{j\Phi(f_x, f_y)}, \qquad 0 \le \alpha \le 2. \tag{7-12}$$

For small α, especially $\alpha = 0$, high-pass filtering results.

Another filter, call pruning [77], is designed to eliminate noise. This filter is a low-pass filter. By prespecifying a threshold T we obtain the filtered image as

$$G'(f_x, f_y) = \begin{cases} \left(|G(f_x, f_y)| - T\right) e^{j\Phi(f_x, f_y)}, & \text{if } |G(f_x, f_y)| \ge T; \\ 0, & \text{if } |G(f_x, f_y)| < T. \end{cases} \tag{7-13}$$

Since the high frequency has less energy, the above filtering will remove the high-frequency portion.

Most of these techniques yield successful results. However, unless a special FFT processor is available, the computation is laborious and requires a huge storage. Note that these techniques require the specification of some constants, e.g., α, f_0; trial runs will make them even more time consuming.

Signal filtering has become a well-developed field [74, 75]. Many of its techniques can be expanded to process two-dimensional images. However, some of them are complex, and some require the specification of a cutoff frequency f_0 or threshold T

8. CONCLUSIONS

In this paper, we survey existing image enhancement techniques. These techniques are classified into four categories: spatial smoothing, gray level rescaling, edge enhancement, and frequency-domain filtering. These procedures are reviewed according to their goals, limitations, practical applicability, and computational complexity. Most techniques are problem oriented. Therefore, we cannot provide general selection criteria. The reader must make his own choice on the basis of the problem at hand, his own judgment, and the available computational facilities.

ACKNOWLEDGMENTS

The authors express their appreciation for the secretarial assistance of Ms. Patricia Halaburka.

REFERENCES

1. K. Preston, Computer processing of biomedical images, *Computer* **9**, 1976, 54.
2. D. Ballard and J. Sklansky, Tumor detection in radiographs, *Comput. Biomed. Res.* **6**, 1973, 299.
3. D. E. Troxel and C. Lynn, Enhancement of news photos, *Computer Graphics Image Processing* **7**, 1978, 266.
4. E. C. Levinthal, W. B. Green, J. A. Cutts, E. D. Jahelka, R. A. Johansen, M. J. Sander, J. B. Seidman, A. T. Young, and L. A. Soderblom, Mariner 9—image processing and products, *Icarus* **18**, 1973, 75.
5. M. J. McDonnell, Restoration of Voyager I images of Io, *Computer Graphics Image Processing* **15**, 1981, 79.
6. D. C. C. Wang, A. H. Vagnucci, and C. C. Li, Image enhancement by gradient inverse weighted smoothing scheme, *Computer Graphics Image Processing* **15**, 1981, 167.
7. B. R. Hunt, The application of constrained least squares estimation to image restoration by digital computer, *IEEE Trans. Comput.* **C-22**, 1973, 805.
8. R. C. Gonzalea and P. Wintz, *Digital Image Processing*, (Applied Mathematics and Computation Series, no. 13), Addison–Wesley, Reading, Mass., 1977.
9. J. M. Prewitt, Object enhancement and extraction, in *Picture Processing and Psychopictorics*, (B. S. Lipkin and A. Rosenfeld, Eds.), p. 70, Academic Press, New York, 1970.
10. A. Rosenfeld and A. C. Kak, *Digital Picture Processing* (Computer Science and Applied Mathematics Series), Academic Press, New York, 1976.
11. H. C. Andrews, Digital image processing, *IEEE Spectrum*, April 1979, 38.
12. H. C. Andrews, Digital image restoration: A survey, *Computer* 7 (5), 1974, 36.
13. T. S. Huang, Some notes on film grain noise, in Woods Hole Summer Study Report on *Restoration of Atmospherically Degraded Images*, Vol. 2, Alexandria, Va., Defense Documentation Center, July 1966.
14. H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice-Hall, New Jersey, 1977.
15. W. K. Pratt, *Digital Image Processing*, Wiley, New York, 1978.
16. J. C. Dainty and R. Shaw, *Image Science*, Academic Press, New York, 1974.
17. J. W. Goodman, Some fundamental properties of speckles, *J. Opt. Soc. Amer.* **66** (11) 1976, 1145.
18. A. Rosenfeld, *Picture Processing by Computer*, Academic Press, New York, 1969.

19. L. S. Davis, A. Rosenfeld, and J. Weszka, Region extraction by averaging and thresholding, *IEEE Trans. Syst. Man Cybern.* **SMC-5**, 1975, 383.
20. R. E. Graham, Snow removal—A noise-stripping process for picture signals, *IRE Trans. Inf. Theor.* **IT-8**, 1962, 129.
21. D. W. Brown, Digital computer analysis and display of the radionuclide scan, *J. Nucl. Med.* **7**, 1966, 740.
22. A. Lev, S. W. Zucker, and A. Rosenfeld, Iterative enhancement of noisy image, *IEEE Trans. Syst. Man Cybern.* **SMC-7**, 1977, 435.
23. M. Kuwahara, K. Hachimura, S. Eiho, and M. Kinoshita, Processing of RI-angiocardiographic images, in *Digital Processing of Biomedical Images*, (K. Preston and M. Onoe, Eds.), Plenum, New York, 1976.
24. F. Tomita and S. Tsuji, Extraction of multiple regions by smoothing in selected neighborhoods, *IEEE Trans. Syst. Man Cybern.* **SMC-7**, 1977, 107.
25. M. Nagao and T. Matsuyama, Edge preserving smoothing, *Computer Graphics Image Processing* **9**, 1979, 394.
26. G. L. Anderson and A. N. Netravali, Image restoration based on a subjective criterion, *IEEE Trans. Syst. Man Cybern.* **SMC-6**, 1976, 845.
27. J. G. Trussell, A fast algorithm for noise smoothing based on a subjective criterion, *IEEE Trans. Syst. Man Cybern.* **SMC-7**, 1977, 677.
28. J. W. Tukey, *Exploratory Data Analysis*, Addison–Wesley, Reading, Mass., 1977.
29. N. S. Jayant, Average and median-based smoothing techniques for improving digital speech quality in the presence of transmission errors, *IEEE Trans. Commun.* **COM-24**, 1976, 1043.
30. B. R. Frieden, A new restoring algorithm for the preferential enhancement of edge gradients, *J. Opt. Soc. Amer.* **66**, 1976, 280.
31. T. S. Huang, G. J. Yang, and G. Y. Tang, A fast two-dimensional median filtering algorithm, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-27**, 1979, 13.
32. P. Narendra, A separable median filter for image noise smoothing, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 1981, 20.
33. R. Kohler and H. Howell, Photographic image enhancement by superimposition of multiple images, *Photo. Sci. Eng.* **7**, 1963, 241.
34. R. Nathan, Picture enhancement for the moon, Mars and man, in *Pictorial Pattern Recognition*, (G. C. Cheng *et al.*, Eds.), p. 239, Thompson, Washington, D.C., 1968.
35. F. Billingley, Applications of digital image processing, *Appl. Opt.* **9**, 1970, 289.
36. E. Hall, Almost uniform distribution for computer image enhancement, *IEEE Trans. Comput.* **C-23**, 1974, 207.
37. H. C. Andrews, A. G. Tescher, and R. P. Kruger, Image processing by digital computer, *IEEE Spectrum* **9** (7) July 1972, 20.
38. E. B. Troy, E. S. Deutsch, and A. Rosenfeld, Gray-level manipulation experiments for texture analysis, *IEEE Trans. Syst. Man Cybern.* **SMC-3**, 1973, 91.
39. E. L. Hall, R. P. Kruger, S. J. Dwyer, D. L. Hall, R. W. McLaren, and G. S. Lodwick, A survey of preprocessing and feature extraction techniques for radiographic images, *IEEE Trans. Comput.* **C-20**, 1971, 103.
40. R. M. Haralick, K. Shanmugan, and I. Dinstein, Texture features for image classification, *IEEE Trans. Syst. Man Cybern.* **SMC-3**, 1973, 610.
41. D. J. Ketcham, R. W. Lowe, and J. W. Weber, Image enhancement techniques for cockpit display, TR-P74-530R, Display System Laboratory, Hughes Aircraft Co., Culver City, Calif., 1974.
42. R. Hummel, Histogram modification techniques, *Computer Graphics Image Processing* **4**, 1975, 209.
43. R. Hummel, Image enhancement by histogram transformation, *Computer Graphics Image Processing* **6**, 1977, 184.
44. W. Frei, Image enhancement by histogram hyperbolization, *Computer Graphics Image Processing* **6**, 1977, 286.
45. G. Wyszecki and W. S. Stiles, *Color Science*, Wiley, New York, 1967.
46. C. J. Bartleson and E. J. Breneman, Brightness perception in complex fields, *J.Opt. Soc. Amer.* **57**, 1967, 953.
47. T. J. Stockham, Image processing in the context of a visual model, *Proc. IEEE* **60**, 1972, 828.
48. A. Rosenfeld and L. S. Davis, Iterative histogram modification, *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 1978, 300.

49. S. Peleg, Iterative histogram modification, 2, *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 1978, 555.
50. H. J. Zweig, E. B. Barrett, and P. C. Hu, Noise-cheating image enhancement, *J.Opt. Soc. Amer.* **65**, 1975, 1347.
51. F. Nadeni and A. A. Sawchuk, Detection of low-contrast image in film-grain noise, *Appl. Opt.* **17**, 1978, 2883.
52. E. E. Christensen, T. S. Curry, and J. Nunnally, *An Introduction to the Physics of Diagnostic Radiology*, Lee & Febiger, Philadelphia, 1972.
53. M. Mojab, L. Garcia, and G. Talge, A new subtraction technique using duplicating film as a final print, *Amer. J. Roentgenol.* **129**, 1977, 528.
54. R. A. Kruger, C. A. Mistretta, A. B. Crummy *et al.*, Digital *K*-Edge subtraction radiography, *Radiology* **125**, 1973, 243.
55. M. G. Ort, E. C. Gregg, and B. Kaufman, Subtraction radiography: Techniques and limitations, *Radiology* **124** (3), 1977, 65.
56. L. S. G. Kovasznay and H. M. Joseph, Image processing, *Proc. IRE* **43**, 1955, 560.
57. L. G. Roberts, Machine perception of three-dimensional solids, in *Optical and Electro-Optical Information Processing*, (J. T. Tippett *et al.*, Eds.), MIT Press, Cambridge, Mass., 1965.
58. A. Arcese, P. Mengert, and W. E. Trombini, Image detection through bipolar correction, *IEEE Trans. Inf. Theor.* **IT-16**, 1970, 534.
59. T. Nagai and T. A. Linuma, A comparison of differential and integral scans, *J. Nucl. Med.* **9**, 1968, 202.
60. A. Rosenfeld and J. Weszka, Picture recognition and scene analysis, *Computer* **9** (5), May 1976, 28.
61. I. E. Abdou and W. K. Pratt, Quantitative design and evaluation of enhancement/thresholding edge detectors, *Proc. IEEE* **67**, 1979, 753.
62. D. Gabor, Information theory in electron microscopy, *Lab. Invest.* **14**, 1965, 801.
63. R. H. Wallis, An approach for the space variant restoration and enhancement of images, *Proceedings, Symposium on Current Math. Prob. in Image Science*, Monterey, Calif., November 1976.
64. J. S. Lee, Digital image enhancement and noise filtering by use of local statistics, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980, 165.
65. J. S. Lee, Refined filtering of image noise using local statistics, *Computer Graphics Image Processing* **15**, 1981, 380.
66. S. W. Levine and H. Mate, Selected electronic techniques for image enhancement, Paper II in *Proceedings, Image Enhancement Seminar*, Soc. Photo. Instr. Engrs., Redondo Beach, Calif. 1963.
67. D. R. Craig, Disenhancement—A negative approach to a positive problem, Paper V in *Proceedings, Image Enhancement Seminar*, Soc. Photo. Instr. Engrs., Redondo Beach, Calif., 1963.
68. A. J. Hannum, Techniques for electronic image enhancement, Paper VII in *Proceedings, Image Enhancement Seminar*, Soc. Photo. Instr. Engrs., Redondo Beach, Calif., 1963.
69. W. F. Schreiber, Wirephoto quality improvement by unsharp masking, *Pattern Recog.* **2**, 1970, 171.
70. J. W. Cooley and T. W. Tukey, An algorithm for machine calculation of complex Fourier series, *Math. Comp.* **19**, 1965, 297.
71. R. C. Singleton, On computing the fast Fourier transform, *Commun. Assoc. Comput. Mach.* **10**, 1967, 647.
72. R. C. Singleton, An Algol procedure for the fast Fourier transform with arbitrary factors, Algorithm 339, *Commun. Assoc. Comput. Mach.* **11**, 1968, 776.
73. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
74. A. V. Oppenheim and R. W. Schaefer, *Digital Signal Processing*, Prentice Hall, Englewood Cliffs, N. J., 1975.
75. T. S. Huang (Ed.), *Picture Processing and Digital Filtering*, Springer, New York, 1975.
76. H. C. Andrews, Monochrome digital image enhancement, *Appl. Opt.* **15**, 1976, 495.
77. P. P. Varoutas, L. R. Nardizzi, and E. M. Stokely, Digital image processing applied to scintillation images from biomedical systems, *IEEE Trans. Biomed. Eng.* **BME-24**, 1977, 337.
78. S. J. Riederer and R. A. Kruger, Intravenous digital subtraction: A summary of recent developments, *Radiology* **147**, 1983, 633.

Exhibit J

Feature enhancement of film mammograms using fixed and adaptive neighborhoods

Richard Gordon and Rangaraj M. Rangayyan

Digital techniques are presented for xerographylike enhancement of features in film mammograms. The mammographic image is first digitized using a procedure for gray scale dynamic range expansion. A pixel operator is then applied to the image, which performs contrast enhancement according to a specified function. The final transformation leads to either a positive or negative mode display as desired. We also present an adaptive neighborhood feature enhancement technique that enhances visibility of objects and details in an image. The availability of the enhanced images should aid diagnosis of breast cancer without requiring additional x-ray dose such as for xeromammography.

I. Introduction

While arguments regarding the relative merits and demerits of film-screen mammography and xeromammography continue,[1–4] it is generally conceded that the availability of both would greatly improve accuracy of diagnosis. However, xeromammography, while depicting clinical details more clearly with better contrast, also requires a much heavier radiation dose than modern film-screen methods (typically five times). A further disadvantage is its lower spatial resolution than film. We present here digital techniques for xerographylike contrast and feature enhancement of film mammograms, which should aid diagnosis without requiring the additional x-ray dose for xeromammography.

Our image processing operation begins with acquisition of a digital image of the mammogram covering the wide range of gray levels present. Since video systems used for image acquisition and digitization have a limited dynamic range, we use two images acquired at different illumination levels to generate a composite digital image, see Ref. 5 for details. A pixel operator is applied to the image which performs contrast enhancement according to a specified function. While implementation of the xerographic enhancement process is possible (see Refs. 6–9 for a description of the modulation transfer function of xerography; see also Refs. 10 and 11 for other techniques for enhancement of mammograms), our method has the flexibility of being able to perform any specified contrast enhancement operation. This feature enables experimentation with a variety of enhancement functions. The final step is transformation of the pixel values for display either in the positive or negative mode.

Contrast enhancement methods depend on differences between neighboring pixels. However, the values of these differences clearly depend on the magnification used. Thus the neighborhood size should be adjusted to the physical resolution. Contrast enhancement also depends on the sharpness of the gradient of density, which depends on the size of the objects and details present in the image. The same object digitized on a coarse raster would have a larger gradient than when digitized on a fine raster. Thus if an image with a wide range of gradients is to be enhanced, we have to use either a varying threshold for enhancement or change the neighborhood size. The former is undesirable, because noise will cause many spurious enhancement operations. The use of a larger neighborhood permits averaging over a number of pixels before differences are calculated, reducing spurious enhancements. We have therefore experimented with methods for adapting the neighborhood size to the local details and features on a pixel-by-pixel basis. The resulting images display features better than those obtained by fixed neighborhood operations. The method is applicable to all types of image and forms the basis for a new, general method for digital image processing.[12,13]

We believe our methods will aid better presentation of details available in conventional film mammograms. The availability of a xerographic representation of the same film mammogram with contrast enhancement (at no extra radiation) and the adaptive neighborhood image should be of considerable diagnostic value.

Both authors are with University of Manitoba, Winnipeg, Manitoba R3E 0W3; R. Gordon is in the Department of Radiology and R. M. Rangayyan is in the Department of Electrical Engineering.

Received 12 April 1983.
0003-6935/84/040560-05$02.00/0.

II. Methods

A. Image Acquisition

We use a Hamamatsu (Waltham, Mass.) C1000 TV camera with a Chalnicon tube to acquire video images of mammograms. The mammogram to be digitized is placed on an illuminator with intensity control. The video signal from the camera is digitized and stored in a Grinnell (San Jose, Calif.) GMR27 frame buffer. Our system has the capability of acquiring a 512 × 480 pixel image with 6-bit resolution (64 gray levels) The digitized image is next transferred to a Control Data Canada (Toronto) Cyber 171 computer via a Camac crate (Kinetic Systems, Lockport, Ill.). Image processing operations are performed by computer programs written in PASCAL (U. Minnesota) on the Cyber and the results transmitted back to the Grinnell system for display. We use a Conrac (Covina, Calif.) color TV monitor for displaying the images, and a Videoslide 35 camera unit (Lang Systems, Menlo Park, Calif.) to take photographs of the images.

As the video system has a limited gray level dynamic range, we acquire two digital images of the same mammogram at low and high illumination levels. The two images are then combined to generate a composite digital image having a wider gray level range; see Ref. 5 for details. By including a calibrated gray scale in the images and transforming pixel values to true optical densities, we also compensate for nonlinearities of the video system. Input images to the enhancement procedures may have pixel values in the optical density range of 0.05 (for white) to 3.04 (for black) as given by the calibrated gray scale used.

B. Contrast Enhancement

A pixel operator is next applied to the image for contrast enhancement. We define the local contrast between a pixel p and the average a of its eight neighbors in the 3 × 3 pixel matrix centered at p as

$$C = |p - a|/(p + a).$$

This definition gives the contrast measure C a range of 0–1. The contrast value is now enhanced according a specified function to a new value C'. A simple enhancement function is $C' = \sqrt{C}$, which increases lower values of contrast by a greater extent and preserves the range of 0–1 for C'. Our contrast enhancement procedure has the advantage over other standard neighborhood operations[14] that a straightforward implementation a any given enhancement function is possible. The new pixel value is then computed from C' and a as follows:

$$p' = a(1 + C')/(1 - C') \text{ if } p \geq a,$$

$$p' = a(1 - C')/(1 + C') \text{ if } p < a.$$

The final step is the transformation of the pixel values to the display range of 0–255. Here we have the choice of specifying either the positive or negative mode for display. If max and min are the maximum and minimum pixel values (optical density) in the image, we have the display pixel value p'' given by

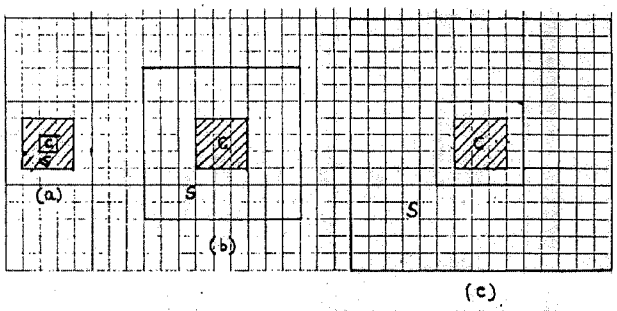


Fig. 1. For a dark square of size 3 × 3 pixels on a light background the contrast is zero in case (a) with center/surround of (1 × 1, 3 × 3); the contrast is maximum in (b) with center/surround of (3 × 3, 9 × 9); the contrast starts to fall with the next step of center/surround of (5 × 5, 15 × 15) as in (c). The center and surround regions are marked c and s in the figures, with the dark square hatched.

$$p'' = 255(p' - \min)/(\max - \min) \quad \text{positive mode,}$$

$$p'' = 255(\max - p')/(\max - \min) \quad \text{negative mode,}$$

with 0 for black and 255 for white in the display. Density windowing can also be incorporated at this stage for better visualization of selected density ranges.

C. Adaptive Neighborhood Processing

The standard neighborhood used in image processing (as above) has a center of one pixel and a surround of eight pixels,[14] if we make an analogy with visual receptive fields in vertebrates. We have found it desirable to retain this ratio of areas for adaptive neighborhood operations as well, and so consider neighborhoods with centers of size $m \times m$ and surrounds of size $3m \times 3m$, where m is any odd integer.[12] The center sizes are kept odd so that the center/surround regions can be centered about a pixel. Thus our neighborhood sizes are (1 × 1, 3 × 3), (3 × 3, 9 × 9), (5 × 5, 15 × 15), etc. The same definition for contrast as before is used, with p as the average of all pixels in the center region, and a as the average of all pixels in the surround region.

To determine the optimum neighborhood size to be used around a given pixel, we thought that a function giving contrast vs the size of the central region would be helpful. As an example, the contrast of a dark square on a light background will be zero when both the center and surround regions fall within the dark square. The contrast increases when the surround region takes in the background and peaks when the center region exactly covers the dark square. Further increase in size of the center region would reduce the contrast, see Fig. 1. We have thus used the first maximum of the contrast vs center size function as the indicator for optimum neighborhood size to be used at a given pixel location. The procedure for contrast enhancement is then as follows: At each pixel the dimensions of the center/surround regions that correspond to the first peak in the contrast are determined. The contrast value for the optimum size chosen is then increased by the square root function. The central pixel is then given a new value determined by the average intensity in the sur-

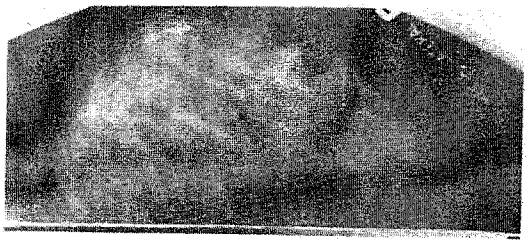


Fig. 2. Digitized mammogram.

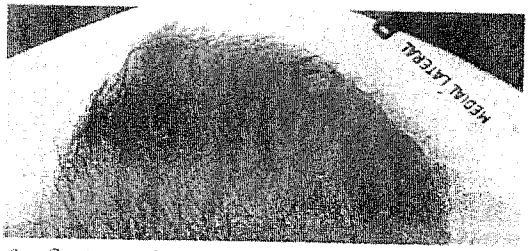


Fig. 3. Contrast enhancement using a fixed 3 × 3 neighborhood.

round region and the new contrast value as before. This is performed at all pixels in the image. The pixel values are then transformed to the display range.

III. Results and Discussion

Figure 2 shows a digitized film mammogram with expanded gray scale dynamic range (see Ref. 5 for details). (The spatial resolution, however, is poor due to digitization with an inadequate matrix of 461 × 269 points, which was the best possible with our equipment.) The corresponding contrast enhanced image using fixed neighborhoods is given in Fig. 3. As is evident, the enhanced image has better contrast and visibility of details. While the square root function has been used here for contrast enhancement, other functions could as well be tried depending on input image characteristics and desired enhancement.

Figure 4 shows a cluster of microcalcifications from a mammogram, enlarged and digitized using a macro-lens on the TV camera. (Dynamic range expansion was not required in this case due to the small area.) Application of the standard edge enhancement technique of adding the digital derivative (see Refs. 10 and 14) yielded Fig. 5, which shows marginal improvement. Figure 6 is the result of our contrast enhancement procedure with a fixed neighborhood size of 3 × 3 (the image is retained in the negative mode). Application of the new adaptive neighborhood feature enhancement procedure resulted in Fig. 7, which has better detail visibility than Figs. 4–6. Note that the uneven background in Fig. 7 is not noise but an enhancement of the weak background pattern in Fig. 4. As can be seen by comparing these figures, the adaptive neighborhood operation enhances objects and features as a whole, rather than enhancing edges only. A pictorial representation of the neighborhood size used is given in Fig. 8, and by comparing this range image with Fig. 4, we see

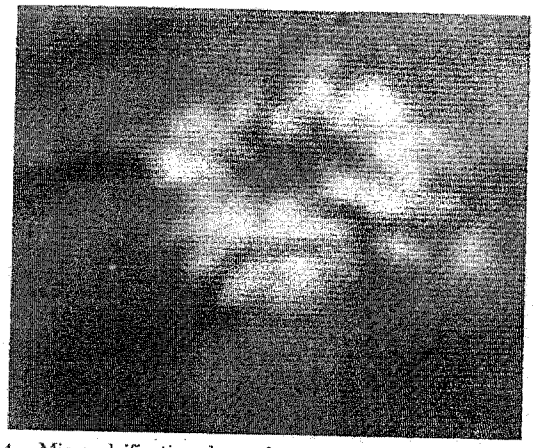


Fig. 4. Microcalcification cluster from a mammogram (enlarged and digitized).

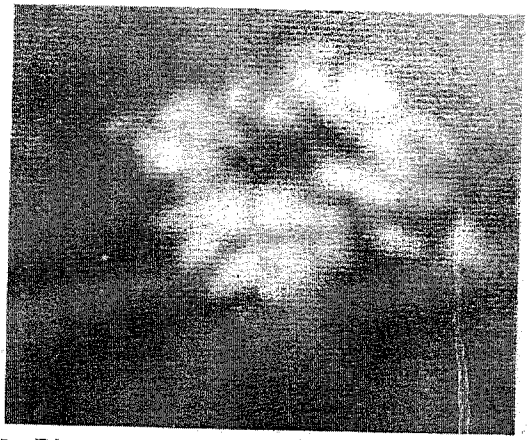


Fig. 5. Edge enhancement by adding digital derivative, achieved by convolving with the 3 × 3 mask (−0.125, −0.125, −0.125, −0.125, 2, −0.125; −0.125, −0.125, −0.125), followed by clipping to the display range of 0–63.

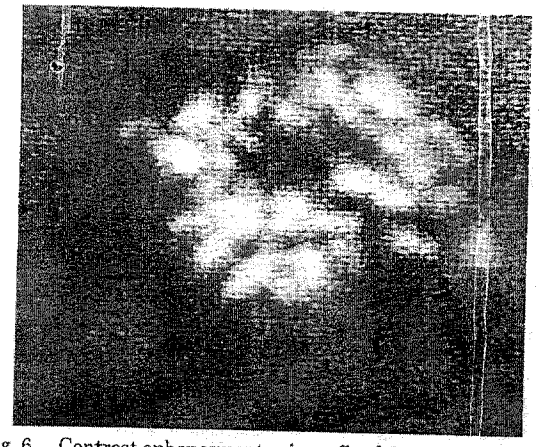


Fig. 6. Contrast enhancement using a fixed 3 × 3 neighborhood.

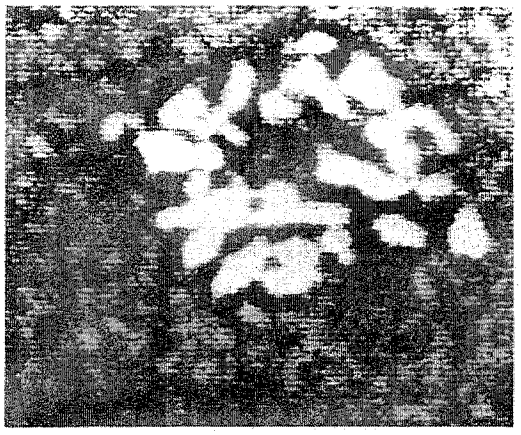


Fig. 7. Adaptive neighborhood feature enhancement.

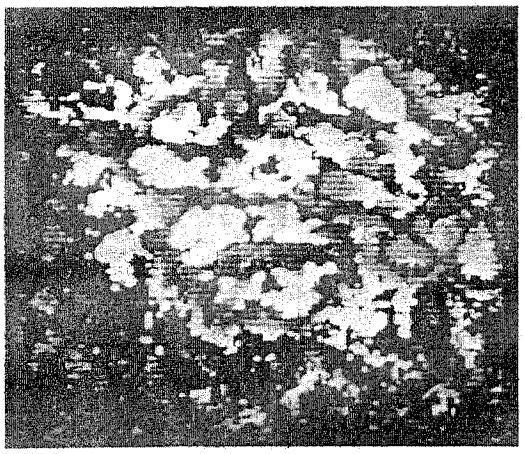


Fig. 8. Picture of neighborhood size used at different pixel locations. Whiter pixels in this range image represent pixels in the original image with larger adaptive neighborhoods.

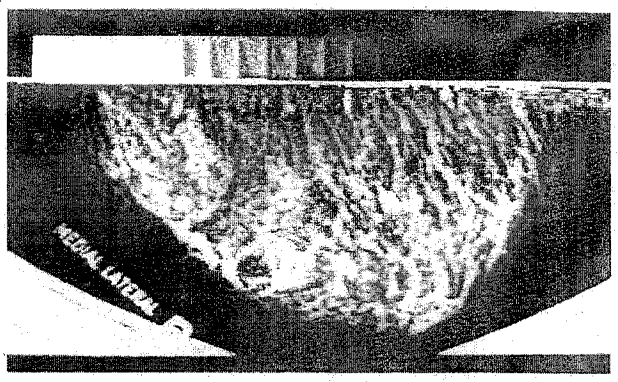


Fig. 9. Adaptive neighborhood enhancement of mammogram in Fig. 2.

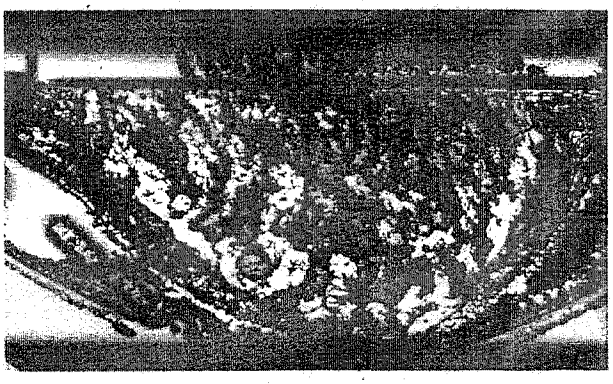


Fig. 10. Range image for Fig. 2.

that the center size used at a pixel location correlates with the size of the object at that location. The images were limited in size to 150 × 150 pixels by the central memory available in our computer. Whole mammograms could be processed with larger computers or special purpose image processing systems.

While MTF (modulation transfer function) analysis of xerography and its simulations has been reported (e.g. Ref. 9), such a linear system theoretical analysis is, strictly speaking, not valid due to the nonlinear nature of the process (see also Ref. 11). Figures 9(a)–(c) give the Fourier transform magnitude spectra for the images in Figs. 4, 6, and 7. It is seen that the two enhancement procedures emphasize higher frequency components, corresponding to edge and contrast enhancement. Note that very small values toward higher frequencies are not seen in Fig. 9(a) due to transformation to the display range of 0–255 (integers). However, since all three spectra have been transformed to the same display range, they can be compared by visual analysis without much error. While we have not performed MTF analysis of the enhancement procedures, the high frequency emphasis nature of the processes is evident from the spectra shown.

Here we have presented new methods for contrast and feature enhancement which enhance the visibility of objects and details in an image. As with all contrast enhancement procedures, the gray scale range of the details is reduced. This may be viewed as an optimal mapping of the wide dynamic range of details present in the image to the limited gray scale range of the human visual system and of the display devices. The contrast enhancement function (the square-root function here) determines the dynamic range of the result: a less severe function could be used if a wider dynamic range is desired. We intend to replace the use of the first peak in the contrast function as an indicator of the optimum neighborhood size by a better scheme which will take into account noise in the function, as well as interactions between objects (ringing). We are now extending the same principles to other image processing techniques, generalizing the concept of pixel-independent image processing.[12]

Judging from our results, as well as those obtained by others, we believe strongly that xeromammography

should not be used at all. The increased dose required, for xeromammography by itself or as an addition to film mammography as commonly practiced, does not justify the enhanced features: digital image processing techniques applied to film mammograms, which have all details recorded due to the wider gray scale dynamic range, can readily provide similar results without additional x-ray dose. Enhancement techniques such as ours can also be viewed as redundancy reduction procedures from an information theoretical point of view.[13] A large scale clinical trial of the techniques with mammograms is contemplated.

Note added in proof: We have rewritten our program to use packed arrays enabling processing of larger images. The adaptive neighborhood enhancement of the mammogram in Fig. 2 is given in Fig. 9, with the corresponding range image in Fig. 10.

We thank John Dunning, Bill Paley, and the late Ronald Laramee of the Computer Department for Health Sciences, University of Manitoba Medical College, for providing image processing hardware and software support. We thank Douglas MacEwan, Department of Radiology, University of Manitoba, and Harold Standing and George Hardy of the Manitoba Mammography Unit, and David Harries, Chief Radiologist at Brandon General Hospital, for cooperation and assistance rendered. Thanks are due also to Aaron Stein and Curtis Quinn for developing software for display of images on the EMI Independent Viewing Station at the Health Sciences Center.

This work was supported by grants from the Manitoba Medical Services Foundation, the Natural Sciences and Engineering Research Council of Canada, and Control Data Corp. This paper was presented in part as a scientific exhibit at the Forty-Fifth Annual Meeting of the Canadian Association of Radiologists, Winnipeg, 29 May–3 June 1982. Richard Gordon also holds appointments in the Departments of Radiology, Electrical Engineering, and Zoology.

References

1. J. J. Pagani, L. W. Bassett, R. H. Gold, J. Benedetti, R. D. Arndt, J. Linsman, and R. L. Scanlan, Am. J. Roentgenol. **135**, 141 (1980.
2. L. Stanton, T. Villafana, J. L. Day, D. A. Lightfoot, and R. E. Stanton, Radiology **132**, 455 (1979).
3. L. N. Rothenburg, R. L. A. Kirch, and R. E. Snyder, Radiology **117**, 701 (1975).
4. J. N. Wolfe, R. P. Dooley, and L. E. Harkins, Cancer **28**, 1569 (1971).
5. R. M. Rangayyan and R. Gordon, "Expanding the Dynamic Range of X-Ray Videodensitometry for Digital Mammography and Teleradiology," Appl. Opt. **00**, 000 (1983), submitted.
6. G. H. Zeman, G. U. V. Rao, and F. A. Osterman, Radiology **119**, 689 (1976).
7. J. W. Boag, Phys. Med. Biol. **18**, 3 (1973).
8. H. E. J. Neugebauer, Appl. Opt. **4**, 453 (1965).
9. R. A. Kilgore, E. C. Gregg, and P. S. Rao, Opt. Eng. **13**, 130 (1974).
10. M. B. McSweeney, P. Sprawls, and R. L. Egan, Am. J. Roentgenol. **140**, 9 (1983).
11. H. J. Trussell, Proc. IEEE **69**, 615 (1981).
12. R. Gordon and R. M. Rangayyan, "Pixel-Independent Image Processing," in preparation.
13. R. Gordon and R. M. Rangayyan, "Radiographic Feature Enhancement, Information Content and Dose Reduction in Mammography and Cardiac Angiography," in *Proceedings, Fifth Annual Conference on Frontiers of Engineering and Computing in Health Care* (IEEE-EMBS), Columbus, Sept. 1983, pp. 161–165.
14. A. Rosenfeld and A. C. Kak, *Digital Picture Processing* (Academic, New York, 1982).

LASER INSTITUTE OF AMERICA

April 2-6, 1984: Laser Safety: Hazard, Inspection, and Control; San Antonio, TX. Provides up-to-date information on Federal, State, and International Laser Safety regulations. Course Fee: $700. Contact: Education Director, Laser Institute of America, 5151 Monroe St., Ste. 118W, Toledo, OH 43623. (419) 882-8706.